State Discrimination With Post-Measurement Information

Manuel A. Ballester, Stephanie Wehner, and Andreas Winter

***Abstract*—We introduce a new state discrimination problem in which we are given additional information about the state after the measurement, or more generally, after a quantum memory bound applies. The following special case plays an important role in quantum cryptographic protocols in the bounded storage model: Given a string x encoded in an unknown basis chosen from a set of mutually unbiased bases (MUBs), you may perform any measurement, but then store at most q qubits of quantum information, and an unlimited amount of classical information. Later on, you learn which basis was used. How well can you compute a function $f(x)$ of x, given the initial measurement outcome, the q qubits, and the additional basis information? We first show a lower bound on the success probability for any balanced function, and any number of mutually unbiased bases, beating the naive strategy of simply guessing the basis. We then show that for two bases, any Boolean function $f(x)$ can be computed perfectly if you are allowed to store just a single qubit, independent of the number of possible input strings x. However, we show how to construct *three* bases, such that you need to store *all* qubits in order to compute $f(x)$ perfectly. We then investigate how much advantage the additional basis information can give for a Boolean function. To this end, we prove optimal bounds for the success probability for the AND and the XOR function for up to three mutually unbiased bases. Our result shows that the gap in success probability can be maximal: *without* the basis information, you can never do better than guessing the basis, but *with* this information, you can compute $f(x)$ perfectly. We also give an example where the extra information does not give any advantage at all.**

***Index Terms*—Bounded quantum storage, quantum cryptography, state discrimination.**

I. Introduction

STATE discrimination with post-measurement information concerns the following task: Consider an ensemble of quantum states, $\mathcal{E} = \{p_{yb}, \rho_{yb}\}$, with double indices $yb \in \mathcal{Y} \times \mathcal{B}$, and a number $q \geq 0$. Suppose Alice sends Bob the state ρ_{yb}, where she alone knows indices y and b. Bob can perform any measurement on his system, but then store at most q qubits (i.e., a Hilbert space of dimension 2^q). Afterwards, Alice tells him b. Bob's goal is now to approximate y as accurately as possible. Here, this means that he has to make a guess $\hat{Y}$, maximizing the success probability

$$p_{\text{succ}} = \sum_{yb} p_{yb} \Pr\{\hat{Y} = y | \text{ state } \rho_{yb}\}.$$

For $|\mathcal{B}| = 1$, i.e., no available post-measurement information, q is irrelevant and Bob's task is to discriminate among states ρ_y, a problem studied since the early days of quantum information science [22]. On the other hand, if the ρ_{yb} all commute, the fact that b comes later—and also the magnitude of q—plays no role as Bob can always measure in the common eigenbasis of the states without losing any information.

Hence, a particular case of the general problem that isolates the aspect of the timing between measurements and side information is one where for each fixed b, the states ρ_{yb} are mutually orthogonal

$$\forall b \forall y \neq z, \quad \rho_{yb} \perp \rho_{zb}. \tag{1}$$

Then the difficulty for Bob and the nontrivial dependence of his probability of success on q derive from the possibility of noncommuting eigenbases of the sets $\{\rho_{yb}\}$ for different b. While for a given b he can distinguish perfectly between the ρ_{yb}, the quantum mechanical measurement-disturbance principle reduces the success probability if this side information is delayed.

In this paper, we focus for the most part on a special case that is of central importance to existing protocols in the bounded quantum storage model [13]. The security of such protocols rests on the realistic assumption that a dishonest player cannot store more than q qubits for long periods of time. In this model, even bit commitment and oblivious transfer can be implemented securely, which is otherwise known to be impossible [29], [31], [14]. In particular, we are interested in the following question: Consider a function $f: \mathcal{X} \to \mathcal{Y}$ between finite sets, and a set of mutually unbiased bases $\mathcal{B}$ (see Definition II.2), given by unitaries $U_0 = \mathbb{1}, U_1, \ldots, U_{|\mathcal{B}|-1}$ on a Hilbert space with basis $\{|x\rangle : x \in \mathcal{X}\}$. Alice chooses a string x and a basis b where xb is drawn from the distribution $P_{X,B}$, prepares the state $U_b|x\rangle$ and sends it to Bob. When Bob receives the state, he may perform any measurement. Afterwards, however, he can store at most q qubits of quantum information. Later, Alice announces which basis she had chosen. Bob's task is now to predict $y = f(x)$ as accurately as possible. That means that the states in our problem are now given by

$$\rho_{yb} = \sum_{x \in f^{-1}(y)} P_{X|B}(x|b) U_b|x\rangle\langle x|U_b^\dagger.$$

Manuscript received September 6, 2006; revised April 23, 2008. Published August 27, 2008 (projected). The work of M. Ballester and S. Wehner was supported by an NWO vici Grant 2004–2009 and by the EU project QAP (IST-2005-15848). The work of S. Wehner was also supported by the National Science Foundation under Grant PHY-0456720. The work of A. Winter was supported by the U.K. EPSRC's "QIP IRC," the EU project QAP, and by a University of Bristol Research Fellowship. Part of this work was completed while S. Wehner was a Ph.D. student at CWI.

M. A. Ballester is with the Centrum voor Wiskunde en Informatica (CWI), 1098 SJ Amsterdam, The Netherlands.

S. Wehner is with the California Institute of Technology, Institute for Quantum Information, Pasadena CA 91125 USA.

A. Winter is with the Department of Mathematics, University of Bristol, Bristol BS8 1TW, U. K.

Communicated by G. Kramer, Associate Editor for Shannon Theory.

Digital Object Identifier 10.1109/TIT.2008.928276

The only difference from (1) is that now we demand the mutual unbiasedness of the joint eigenbases of the $\{\rho_{yb}\}$ for different b. How well can Bob compute $f(x)$ given the classical outcome of his earlier measurement, the q qubits, and the additional basis information? In the context of cryptographic protocols [13], Bob is a dishonest player who tries to learn some function of the encoded string conditioned on the fact that he will later learn the basis and the function. In the oblivious transfer protocol of [13], Alice uses two mutually unbiased bases, and secretly chooses a function from a set of predetermined functions. She then tells Bob which function he should evaluate together with the basis information b. This makes the protocols more complicated and so one might wonder whether it is possible to use a fixed Boolean function instead, a question which stood at the beginning of the current investigation (see [13, Sec. 3.6] in the `arXiv` version, where the XOR of an even number of bits is shown to be insufficient). However, we show that this is not possible in the suggested protocol. In particular, we show that for two bases and *any* Boolean function f, Bob can succeed with probability at least $1/2 + 1/(2\sqrt{2})$, even if he cannot store any qubits at all. Surprisingly, it also turns out that Bob can determine $f(x)$ perfectly, if he can store just a *single* qubit. We show that one qubit is sufficient no matter how long the input string x actually is. Behind our proof, there is an algebraic framework that allows us, in principle, to determine the minimal quantum memory resources required and the optimal strategy to succeed with probability 1 for any number of bases and any function f. However, it turns out that we can construct *three* bases, such that Bob needs to store *all* qubits in order to compute a Boolean function perfectly.

In general, we also show a lower bound on Bob's optimal success probability for any balanced function $f: \mathcal{X} \to \mathcal{Y}$, and any number of mutually unbiased bases if he cannot store any qubits. Our bound is strictly better than what Bob could achieve by guessing the basis. Our problem also has an interpretation in the light of communication complexity. Suppose Alice is given b, Bob is given the state ρ_{yb}, and Alice cannot herself send any information to Bob. If classical communication is free, what is the minimal number of qubits Bob needs to communicate to Alice such that Alice learns y? It turns out that if there exists a strategy for Bob to compute y in our original task while storing only q qubits, he will also need to send exactly q qubits, and his classical measurement outcome, to allow Alice to learn y: Alice now simply performs the measurement Bob would have done in our original task after he received b.

It is an interesting problem to consider how much the extra basis information helps Bob to compute $f(x)$. To this end, we first examine how well Bob can compute the AND and XOR of x *without* using the additional basis information. We prove optimal bounds for computing the AND and XOR function on a string of length n for two and three mutually unbiased bases. In particular, we show that for two mutually unbiased bases and the XOR function on strings of even length, Bob's probability of success is at most $3/4$, and there exists a strategy which achieves it. This means that his trivial strategy of guessing the basis and taking the measurement outcome in that basis to be the real answer, is optimal. Interestingly, adding the third basis does not change his success probability of $3/4$, whereas intuitively one would expect it to be lower. Surprisingly, for three bases, if we choose a nonuniform prior distribution over the strings of length n, it actually becomes harder for Bob to compute the XOR. We show that there exists a nonuniform distribution such that he can never succeed more than using the trivial strategy of guessing the outcome. No measurement he can perform will give him any more information. We then examine the case that the length of the string n is odd. Here, Bob can succeed only with probability $1/2+1/(2\sqrt{2})$ which is optimal. We prove that for *any* Boolean function f, Bob's probability of success is upper-bounded by $1/2 + 1/(2\sqrt{|\mathcal{B}|})$ if he does not receive any basis information.

We then examine how well Bob can do *with* the additional basis information.

We show that for the XOR function on strings of even length, Bob can now compute the value of the function perfectly and give an explicit measurement strategy for Bob. For two bases, this means that the gap can be maximal: *without* the basis information Bob cannot do better than the trivial strategy of guessing the basis, however, *with* this extra information Bob always succeeds. It also means that the gap can be minimal: For the XOR on strings of odd length the extra information does not help Bob at all. For three bases we obtain the maximum gap only for a nonuniform prior. Finally, we also give an optimal strategy for computing the AND from the given state and the post-measurement information.

A. Related Work

State discrimination itself has received considerable attention in the past: Alice prepares a quantum state drawn from a collection of possible quantum states. Bob's goal is now to determine the identity of the state. The new twist in the present work is that after the measurement, or more generally after a memory bound applies, he is given additional information. For the case of only two (mixed) states, the optimal measurement for traditional minimum-error state discrimination was found by Helstrom [22]. The case of multiple (mixed) states was already considered by Holevo [23] and Yuen, Kennedy, and Lax [37] in the 1970s, and they have given the necessary conditions for a measurement to be optimal. Yuen *et al.* also showed these conditions to be sufficient and demonstrated that the problem of finding the optimal measurement can be expressed as convex optimization problem. Discriminating between multiple mixed states remains a difficult problem and it is usually hard to derive explicit measurements and bounds from these conditions. Optimal measurements are known only for special state sets, which satisfy certain symmetry properties [19], [4], [6], [1], or we consider pure qubit states where the prior distribution over the states is uniform [26].

Many convex optimization problems can be solved using semidefinite programming. Eldar [16] and Eldar, Megretski, and Verghese [18] used semidefinite programming to solve state discrimination problems, which is one of the techniques we will also use here. The square-root measurement [20] (also called pretty good measurement) is an easily constructed measurement to distinguish quantum states, however, it is only optimal for very specific sets of states [17], [19]. Mochon constructed specific pure state discrimination problems for which the square-root measurement is optimal [32]. We will use a variant of the square-root measurement as well. Furthermore,

our problem is related to the task of state filtering [7], [9], [10] and state classification [34]. Here, Bob's goal is to determine whether a given state is either a specific state or one of several other possible states, or, more generally, which subset of states a given state belongs to. Our scenario differs, because we deal with mixed states and Bob is allowed to use post-measurement information. Much more is known about pure state discrimination problems and the case of unambiguous state discrimination where we are not allowed to make an error. Since we concentrate on mixed states, we refer to [8] for an excellent survey on the extended field of state discrimination.

Regarding state discrimination with post-measurement information, special instances of the general problem have occurred in the literature under the heading "mean king's problem" [2], [27], where the stress was on the usefulness of entanglement. Furthermore, it should be noted that prepare-and-measure quantum key distribution schemes of the BB84 type also lead to special cases of this problem: When considering optimal individual attacks, the eavesdropper is faced with the task of extracting maximal information about the raw key bits, encoded in an unknown basis, that she learns later during basis reconciliation.

II. Preliminaries

A. Notation and Tools

We will need the following notions. The Bell basis is given by the vectors $|\Phi^{\pm}\rangle = (|00\rangle \pm |11\rangle)/\sqrt{2}$ and $|\Psi^{\pm}\rangle = (|01\rangle \pm |10\rangle)/\sqrt{2}$. Furthermore, let $f^{-1}(y) = \{x \in \mathcal{X} | f(x) = y\}$. We say that a function f is balanced if and only if any element in the image of f is generated by equally many elements in the pre-image of f, i.e., there exists a $k \in \mathbb{N}$ such that $\forall y \in \mathcal{Y}$: $|f^{-1}(y)| = k$. We also use the notation $[m] = \{1, \ldots, m\}$. $A^{\dagger}$ is the conjugate transpose of matrix A. A *positive semidefinite* $n \times n$ matrix A is a Hermitian matrix such that $x^* A x \geq 0$ for all $x \in \mathbb{C}^n$ [24]. A is said to be *positive definite*, if it is positive semidefinite and $x^* A x = 0$ implies $x = 0$. We use $A \geq 0$ and $A > 0$ to indicate that A is positive semidefinite and positive definite, respectively. Finally, $\|A\|_1 = \mathrm{Tr}\sqrt{A^{\dagger} A}$ is the trace norm. The first tool we use is the following well-known result.

Theorem II.1: (Helstrom [22]). Suppose we are given states ρ_0 with probability q, and ρ_1 with probability $1 - q$. Then the probability to determine whether the state was ρ_0 and ρ_1 is at most

$$p = \frac{1}{2}[1 + ||q\rho_0 - (1-q)\rho_1||_1].$$

The measurement that achieves p is given by M_0, and $M_1 = \mathbb{1} - M_0$, where M_0 is the projector onto the positive eigenspace of $q\rho_0 - (1-q)\rho_1$. □

Second, we will make use of semidefinite programming, which is a special case of convex optimization. We refer to [11] for an in-depth introduction. The goal of semidefinite programming is to solve he following semidefinite program (SDP) in terms of the variable $X \in S^n$

$$\begin{array}{ll} \text{maximize} & \mathrm{Tr}(CX) \\ \text{subject to} & \mathrm{Tr}(A_i X) = b_i, i = 1, \ldots, p \\ & X \geq 0 \end{array}$$

for given matrices $C, A_1, \ldots, A_p \in S^n$ where S^n is the space of symmetric $n \times n$ matrices. X is called *feasible*, if it satisfies all constraints. An important aspect of semidefinite programming is duality. Intuitively, the idea behind Lagrangian duality is to extend the objective function (here $\mathrm{Tr}(CX)$) with a weighted sum of the constraints in such a way, that we will be penalized if the constraints are not fulfilled. The weights then correspond to the dual variables. Optimizing over these weights then gives rise to the *dual problem*. The original problem is called the *primal problem*. Let d' denote the optimal value of the dual problem, and p' the optimal value of the primal problem stated above. Weak duality says that $d' \geq p'$. In particular, if we have $d' = p'$ for a feasible dual and primal solution, respectively, we can conclude that both solutions are optimal.

We will also need the notion of mutually unbiased bases, which was introduced in [36]. The following definition closely follows the one given in [5].

Definition II.2: Let $B_1 = \{|\phi_1\rangle, \ldots, |\phi_d\rangle\}$ and $B_2 = \{|\psi_1\rangle, \ldots, |\psi_d\rangle\}$ be two orthonormal bases in a d-dimensional Hilbert space. They are said to be *mutually unbiased* if and only if $|\langle\phi_i|\psi_j\rangle| = 1/\sqrt{d}$, for every $i, j = 1, \ldots, d$. A set $\{\mathcal{B}_1, \ldots, \mathcal{B}_m\}$ of orthonormal bases in $\mathbb{C}^d$ is called a *set of mutually unbiased bases (MUBs)* if each pair of bases $\mathcal{B}_i$ and $\mathcal{B}_j$ is mutually unbiased.

In any dimension d, the number of MUBs is at most $d + 1$ [5]. Explicit constructions are known if d is a prime power [5], [36] or a square [35]. We say that a set of unitaries $\{U_s\}$ gives rise to $|\{U_s\}|$ MUBs, if those unitaries generate $|\{U_s\}|$ MUBs when applied to the basis vectors of the computational basis.

B. Definitions

We now give a more formal description of our problem. Let $\mathcal{Y}$ and $\mathcal{B}$ be finite sets and let $P_{Y,B} = \{p_{yb}\}$ be a probability distribution over $\mathcal{Y} \times \mathcal{B}$. Consider an ensemble of quantum states $\mathcal{E} = \{p_{yb}, \rho_{yb}\}$. We assume that $\mathcal{Y}$, $\mathcal{B}$, $\mathcal{E}$, and $P_{Y,B}$ are known to both Alice and Bob. Suppose now that Alice chooses $yb \in \mathcal{Y} \times \mathcal{B}$ according to probability distribution $P_{Y,B}$, and sends ρ_{yb} to Bob. We can then define the tasks as follows.

Definition II.3: State discRimination ($\mathrm{STAR}(\mathcal{E})$) is the following task for Bob. Given ρ_{yb}, determine y. He can perform any operation on ρ_{yb} immediately upon receipt.

Definition II.4: State discRimination with Post-measurement Information (PI_q-STAR ($\mathcal{E}$)) is the following task for Bob. Given ρ_{yb}, determine y, where Bob can use the following sources of information in succession.

1) He can perform any measurement on ρ_{yb} immediately upon reception. Afterwards, he can store at most q qubits of quantum information about ρ_{yb}, and an unlimited amount of classical information.
2) After Bob's measurement, Alice announces b.
3) Then, he may measure the remaining q qubits depending on b and the measurement outcome obtained in 1.

We also say that *Bob succeeds at* $\mathrm{STAR}(\mathcal{E})$ or PI_q-$\mathrm{STAR}(\mathcal{E})$ with probability p if and only if p is the average success probability $p = \sum_{yb} p_{yb} \mathrm{Pr}\{\hat{Y} = y | \text{state } \rho_{yb}\}$, where $\mathrm{Pr}\{\hat{Y} =$

$y|$ state $\rho_{yb}\}$ is the probability that Bob correctly determines y given ρ_{yb} in the case of STAR, and in addition using information sources 1, 2 and 3 in the case of PI-STAR.

In this paper, we are interested in the following special case: Consider a function $f:\mathcal{X} \to \mathcal{Y}$ between finite sets, and a set of MUBs $\mathcal{B}$ generated by a set of unitaries $U_0, U_1, \ldots, U_{|\mathcal{B}|-1}$ acting on a Hilbert space with basis $\{|x\rangle : x \in \mathcal{X}\}$. Take $|\Phi_b^x\rangle = U_b|x\rangle$. Let P_X and P_B be probability distributions over $\mathcal{X}$ and $\mathcal{B}$, respectively. We assume that f, $\mathcal{X}$, $\mathcal{Y}$, $\mathcal{B}$, P_X, P_B, and the set of unitaries $\{U_b|b \in \mathcal{B}\}$ are known to both Alice and Bob. Suppose now that Alice chooses $x \in \mathcal{X}$ and $b \in \mathcal{B}$ independently according to probability distributions P_X and P_B, respectively, and sends $|\Phi_b^x\rangle$ to Bob. Bob's goal is now to compute $y = f(x)$. We thus obtain an instance of our problem with states $\rho_{yb} = \sum_{x \in f^{-1}(y)} P_X(x)|\Phi_b^x\rangle\langle\Phi_b^x|$. We write $\mathrm{STAR}(f)$ and $\mathrm{PI}_q\text{-}\mathrm{STAR}(f)$ to denote both problems in this special case. We concentrate on the case of MUBs, as this case is most relevant to our initial goal of analyzing protocols for quantum cryptography in the bounded storage model [13].

Here, we will make use of the basis set $\mathcal{B} = \{+, \times, \odot\}$, where $\mathcal{B}_+ = \{|0\rangle, |1\rangle\}$ is the computational basis,

$$\mathcal{B}_\times = \left\{ \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \right\}$$

is the Hadamard basis, and

$$\mathcal{B}_\odot = \left\{ \frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle), \frac{1}{\sqrt{2}}(|0\rangle - i|1\rangle) \right\}$$

is what we call the K-basis. The unitaries that give rise to these bases are $U_+ = \mathbb{1}$, $U_\times = H$, and $U_\odot = K$ with $K = (\mathbb{1} + i\sigma_x)/\sqrt{2}$, respectively. The Hadamard matrix is given by $H = \frac{1}{\sqrt{2}}(\sigma_x + \sigma_z)$. σ_x, σ_z, and σ_y are the well-known Pauli matrices. We generally assume that Bob has no *a priori* knowledge about the outcome of the function and about the value of b. This means that b is chosen uniformly at random from $\mathcal{B}$, and, in the case of balanced functions, that Alice chooses x uniformly at random from $\mathcal{X}$. More generally, the distribution is uniform on all $f^{-1}(y)$ and such that each value $y \in \mathcal{Y}$ is equally likely.

C. A Trivial Bound: Guessing the Basis

Note that a simple strategy for Bob is to guess the basis, and then measure. This approach leads to a lower bound on the success probability for both STAR and PI-STAR. In short, we have the following.

Lemma II.5: Let $P_X(x) = \frac{1}{2^n}$ for all $x \in \{0,1\}^n$. Let $\mathcal{B}$ denote the set of bases. Then for any balanced function $f:\mathcal{X} \to \mathcal{Y}$ Bob succeeds at $\mathrm{STAR}(f)$ and $\mathrm{PI}_0\text{-}\mathrm{STAR}(f)$ with probability at least

$$p_{\text{guess}} = \frac{1}{|\mathcal{B}|} + \left(1 - \frac{1}{|\mathcal{B}|}\right)\frac{1}{|\mathcal{Y}|}. \qquad \square$$

Our goal is to beat this bound. We show that for PI-STAR, Bob can indeed do much better.

III. No Post-Measurement Information

We first consider the standard case of state discrimination. Here, Alice does not supply Bob with any additional post-measurement information. Instead, Bob's goal is to compute $y = f(x)$ immediately. This analysis will enable us to gain interesting insights into the usefulness of post-measurement information later.

A. Two Simple Examples

We now examine two simple one-qubit examples of a state discrimination problem, which we make use of later on. Here, Bob's goal is to learn the value of a bit which has been encoded in two or three MUBs while he does not know which basis has been used.

Lemma III.1: Let $x \in \{0,1\}$, $P_X(x) = \frac{1}{2}$, and $f(x) = x$. Let $\mathcal{B} = \{+, \times\}$ with $U_+ = \mathbb{1}$ and $U_\times = H$. Then Bob succeeds at $\mathrm{STAR}(f)$ with probability at most

$$p = \frac{1}{2} + \frac{1}{2\sqrt{2}}.$$

There exists a strategy for Bob that achieves p.

Proof: The probability of success follows from Theorem II.1 with $\rho_0 = \frac{1}{2}(|0\rangle\langle 0| + H|0\rangle\langle 0|H)$, $\rho_1 = \frac{1}{2}(|1\rangle\langle 1| + H|1\rangle\langle 1|H)$, and $q = 1/2$. ∎

Lemma III.2: Let $x \in \{0,1\}$, $P_X(x) = \frac{1}{2}$, and $f(x) = x$. Let $\mathcal{B} = \{+, \times, \odot\}$ with $U_+ = \mathbb{1}$, $U_\times = H$, and $U_\odot = K$. Then Bob succeeds at $\mathrm{STAR}(f)$ with probability at most

$$p = \frac{1}{2} + \frac{1}{2\sqrt{3}}.$$

There exists a strategy for Bob that achieves p.

Proof: The proof is identical to that of Lemma III.1 using $\rho_0 = \frac{1}{3}(|0\rangle\langle 0| + H|0\rangle\langle 0|H + K|0\rangle\langle 0|K^\dagger)$, $\rho_1 = \frac{1}{3}(|1\rangle\langle 1| + H|1\rangle\langle 1|H + K|1\rangle\langle 1|K^\dagger)$, and $q = 1/2$. ∎

B. An Upper Bound for All Boolean Functions

We now show that for any Boolean function f and any number of MUBs, the probability that Bob succeeds at $\mathrm{STAR}(f)$ is very limited.

Theorem III.3: Let $|\mathcal{Y}| = 2$, let f be a balanced function, and let $\mathcal{B}$ be a set of MUBs. Then Bob succeeds at $\mathrm{STAR}(f)$ with probability at most

$$p = \frac{1}{2} + \frac{1}{2\sqrt{|\mathcal{B}|}}.$$

In particular, for $|\mathcal{B}| = 2$, we obtain $(1 + 1/\sqrt{2})/2 \approx 0.853$; for $|\mathcal{B}| = 3$, we obtain $(1 + 1/\sqrt{3})/2 \approx 0.789$.

Proof: The probability of success is given by Theorem II.1 where for $y \in \{0,1\}$

$$\rho_y = \frac{1}{2^{n-1}|\mathcal{B}|}\sum_{b=1}^{|\mathcal{B}|} P_{yb}$$

with

$$P_{yb} = \sum_{x \in f^{-1}(y)} U_b |x\rangle\langle x| U_b^\dagger.$$

Using the Cauchy– Schwarz inequality we can show that

$$\begin{aligned} ||\rho_0 - \rho_1||_1^2 &= [\mathrm{Tr}(|\rho_0 - \rho_1|\mathbb{I})]^2 && (2)\\ &\leq \mathrm{Tr}[(\rho_0 - \rho_1)^2]\mathrm{Tr}[\mathbb{I}^2] && (3)\\ &= 2^n \mathrm{Tr}[(\rho_0 - \rho_1)^2] && (4)\end{aligned}$$

or

$$||\rho_0 - \rho_1||_1 \leq \sqrt{2^n \mathrm{Tr}[(\rho_0 - \rho_1)^2]}. \quad (5)$$

A simple calculation shows that

$$\mathrm{Tr}[(\rho_0 - \rho_1)^2] = \frac{4}{2^n |\mathcal{B}|}.$$

The theorem then follows from the previous equation, together with Theorem II.1 and (2). ∎

C. AND *Function*

One of the simplest functions to consider is the AND function. Recall, that we always assume that Bob has no *a priori* knowledge about the outcome of the function. In the case of the AND, this means that we are considering a very specific prior: with probability $1/2$ Alice will choose the only string x for which $\mathrm{AND}(x) = 1$. Without any post-measurement information, Bob can already compute the AND quite well.

Theorem III.4: Let $P_X(x) = \frac{1}{2(2^n-1)}$ for all $x \in \{0,1\}^n \setminus \{1\cdots 1\}$ and $P_X(1\cdots 1) = \frac{1}{2}$. Let $\mathcal{B} = \{+, \times\}$ with $U_+ = \mathbb{I}^{\otimes n}$, $U_\times = H^{\otimes n}$, and $P_B(+) = P_B(\times) = 1/2$. Then Bob succeeds at $\mathrm{STAR}\,(\mathrm{AND})$ with probability at most

$$p = \begin{cases} \frac{1}{2} + \frac{1}{2\sqrt{2}}, & \text{if } n = 1 \\ 1 - \frac{1}{2(2^n-1)}, & \text{if } n \geq 2. \end{cases} \quad (6)$$

There exists a strategy for Bob that achieves p.

Proof: Let $|c_1\rangle = |1\rangle^{\otimes n}$ and $|h_1\rangle = [H|1\rangle]^{\otimes n}$. Equation (6) is obtained by substituting

$$\rho_0 = \frac{1}{2}\left[\frac{\mathbb{I} - |c_1\rangle\langle c_1|}{2^n - 1} + \frac{\mathbb{I} - |h_1\rangle\langle h_1|}{2^n - 1}\right]$$
$$\rho_1 = \frac{|c_1\rangle\langle c_1| + |h_1\rangle\langle h_1|}{2}$$

and $q = 1/2$ in Theorem II.1. ∎

In Theorem IV.3, we will show an optimal bound for the case that Bob does indeed receive the extra information. By comparing the previous equation with (8), one can see that for $n = 1$ announcing the basis does not help. However, for $n > 1$ we will observe an improvement of $[2(2^n + 2^{n/2} - 2)]^{-1}$.

D. XOR *Function*

The XOR function provides an example of a Boolean function where we observe both the largest advantage as well as the smallest advantage in receiving post-measurement information: For strings of even length, we will show that without the extra information Bob can never do better than guessing the basis. For strings of odd length, however, he can do quite a bit better. Interestingly, it will turn out that in this case the post-measurement information is completely useless to him. We first investigate how well Bob does at $\mathrm{STAR}(\mathrm{XOR})$ for two bases.

Theorem III.5: Let $P_X(x) = \frac{1}{2^n}$ for all $x \in \{0,1\}^n$. Let $\mathcal{B} = \{+, \times\}$ with $U_+ = \mathbb{I}^{\otimes n}$, $U_\times = H^{\otimes n}$, and $P_B(+) = P_B(\times) = 1/2$. Then Bob succeeds at $\mathrm{STAR}(\mathrm{XOR})$ with probability at most

$$p = \begin{cases} \frac{3}{4} & \text{if } n \text{ is even} \\ \frac{1}{2}\left(1 + \frac{1}{\sqrt{2}}\right), & \text{if } n \text{ is odd.} \end{cases}$$

There exists a strategy for Bob that achieves p.

Proof: Our proof works by induction on n. The case of $n = 1$ was addressed in Lemma III.1. Now, consider $n = 2$: Let $\sigma_0^2 = \frac{1}{2}(\rho_{0+}^2 + \rho_{0\times}^2)$ and $\sigma_1^2 = \frac{1}{2}(\rho_{1+}^2 + \rho_{1\times}^2)$, where ρ_{0+}^2 and ρ_{1+}^2 are defined as

$$\rho_{o_b b}^n = \frac{1}{2^{n-1}} \sum_{x \in \{0,1\}^n,\, x \in \mathrm{XOR}^{-1}(o_b)} U_b |x\rangle\langle x| U_b^\dagger$$

with $o_b \in \{0,1\}$ and $b \in \mathcal{B}$. We have $||\sigma_0^2 - \sigma_1^2||_1 = 1$.

We now show that the trace distance does not change when we go from strings of length n to strings of length $n+2$: Note that we can express ρ_{yb}^{n+2} as

$$\begin{aligned} \rho_{0+}^{n+2} &= \frac{1}{2}\left(\rho_{0+}^n \otimes \rho_{0+}^2 + \rho_{0+}^n \otimes \rho_{1+}^2\right)\\ \rho_{0\times}^{n+2} &= \frac{1}{2}\left(\rho_{0\times}^n \otimes \rho_{0\times}^2 + \rho_{1\times}^n \otimes \rho_{1\times}^2\right)\\ \rho_{1+}^{n+2} &= \frac{1}{2}\left(\rho_{0+}^n \otimes \rho_{1+}^2 + \rho_{1+}^n \otimes \rho_{0+}^2\right)\\ \rho_{1\times}^{n+2} &= \frac{1}{2}\left(\rho_{0\times}^n \otimes \rho_{1\times}^1 + \rho_{1\times}^n \otimes \rho_{0\times}^2\right). \end{aligned} \quad (7)$$

Let $\sigma_0^n = \frac{1}{2}(\rho_{0+}^n + \rho_{0\times}^n)$ and $\sigma_1^n = \frac{1}{2}(\rho_{1+}^n + \rho_{1\times}^n)$. A small calculation shows that

$$\begin{aligned} \sigma_0^{n+2} - \sigma_1^{n+2} = \frac{1}{8}\big[&\left(\rho_{0+}^n + \rho_{0\times}^n - \rho_{1+}^n - \rho_{1\times}^n\right) \otimes |\Phi^+\rangle\langle\Phi^+|\\ &- \left(\rho_{0+}^n + \rho_{0\times}^n - \rho_{1+}^n - \rho_{1\times}^n\right) \otimes |\Psi^-\rangle\langle\Psi^-|\\ &+ \left(\rho_{0+}^n + \rho_{1\times}^n - \rho_{1+}^n - \rho_{0\times}^n\right) \otimes |\Phi^-\rangle\langle\Phi^-|\\ &- \left(\rho_{0+}^n + \rho_{1\times}^n - \rho_{1+}^n - \rho_{0\times}^n\right) \otimes |\Psi^+\rangle\langle\Psi^+|\big]. \end{aligned}$$

We then get that

$$||\sigma_0^{n+2} - \sigma_1^{n+2}||_1 = \frac{1}{2}\left(||\sigma_0^n - \sigma_1^n||_1 + ||\tilde{\sigma}_0^n - \tilde{\sigma}_1^n||_1\right)$$

where $\tilde{\sigma}_0^n = \frac{1}{2}(\rho_{0+}^n + \rho_{1\times}^n)$ and $\tilde{\sigma}_n^1 = \frac{1}{2}(\rho_{1+}^n + \rho_{0\times}^n)$. Consider the unitary $U = \sigma_x^{\otimes n}$ if n is odd, and $U = \sigma_x^{\otimes n-1} \otimes \mathbb{I}$ if n is even. It is easy to verify that $\sigma_0^n = U\tilde{\sigma}_0^n U^\dagger$ and $\sigma_1^n = U\tilde{\sigma}_1^n U^\dagger$. We thus have that $||\sigma_0^n - \sigma_1^n||_1 = ||\tilde{\sigma}_0^n - \tilde{\sigma}_1^n||_1$ and therefore

$$||\sigma_0^{n+2} - \sigma_1^{n+2}||_1 = ||\sigma_0^n - \sigma_1^n||_1.$$

It then follows from Helstrom's theorem [22] that the maximum probability to distinguish σ_0^{n+2} from σ_1^{n+2} and thus compute the XOR of the $n+2$ bits is given by

$$\frac{1}{2} + \frac{||\sigma_0^n - \sigma_1^n||_1}{4}$$

which gives the claimed result. ∎

A similar argument is possible, if we use three MUBs. Intuitively, one might expect Bob's chance of success to drop as we had more bases. Interestingly, however, we obtain the same bound of $3/4$ if n is even.

Theorem III.6: Let $P_X(x) = \frac{1}{2^n}$ for all $x \in \{0,1\}^n$. Let $\mathcal{B} = \{+, \times, \odot\}$ with $U_+ = \mathbb{I}^{\otimes n}$, $U_\times = H^{\otimes n}$, and $U_\odot = K^{\otimes n}$ with $P_B(+) = P_B(\times) = P_B(\odot) = 1/3$. Then Bob succeeds at $\mathrm{STAR}(\mathrm{XOR})$ with probability at most

$$p = \begin{cases} \frac{3}{4}, & \text{if } n \text{ is even} \\ \frac{1}{2}\left(1 + \frac{1}{\sqrt{3}}\right), & \text{if } n \text{ is odd.} \end{cases}$$

There exists a strategy for Bob that achieves p.

Proof: Our proof is very similar to the case of only 2 MUBs. The case of $n = 1$ follows from Lemma III.2. This time, we have for $n = 2$: $\sigma_0^2 = \frac{1}{3}(\rho_{0+}^2 + \rho_{0\times}^2 + \rho_{0\odot}^2)$ and $\sigma_1^2 = \frac{1}{3}(\rho_{1+}^2 + \rho_{1\times}^2 + \rho_{1\odot}^2)$. We have $||\sigma_0^2 - \sigma_1^2||_1 = 1$. We again show that the trace distance does not change when we go from strings of length n to strings of length $n+2$. We can compute

$$\begin{aligned} \sigma_0^{n+2} - \sigma_1^{n+2} = \frac{1}{4}\big[&(\bar{\sigma}_1^n - \bar{\sigma}_0^n) \otimes |\Phi^+\rangle\langle\Phi^+| \\ &- (\hat{\sigma}_1^n - \hat{\sigma}_0^n) \otimes |\Psi^-\rangle\langle\Psi^-| \\ &+ (\check{\sigma}_1^n - \check{\sigma}_0^n) \otimes |\Phi^-\rangle\langle\Phi^-| \\ &- (\sigma_1^n - \sigma_0^n) \otimes |\Psi^+\rangle\langle\Psi^+|\big] \end{aligned}$$

where $\bar{\sigma}_1^n = (\rho_{0+}^n + \rho_{0\times}^n + \rho_{1\odot}^n)/3$, $\bar{\sigma}_0^n = (\rho_{1+}^n + \rho_{1\times}^n + \rho_{0\odot}^n)/3$, $\hat{\sigma}_1^n = (\rho_{0+}^n + \rho_{1\times}^n + \rho_{0\odot}^n)/3$, $\hat{\sigma}_0^n = (\rho_{1+}^n + \rho_{0\times}^n + \rho_{1\odot}^n)/3$, $\check{\sigma}_0^n = (\rho_{1+}^n + \rho_{0\times}^n + \rho_{0\odot}^n)/3$, and $\check{\sigma}_0^n = (\rho_{0+}^n + \rho_{1\times}^n + \rho_{1\odot}^n)/3$. Consider the unitaries $\bar{U} = \sigma_y^{\otimes n}$, $\hat{U} = \sigma_x^{\otimes n}$, and $\check{U} = \sigma_z^{\otimes n}$ if n is odd, and $\bar{U} = \sigma_y^{\otimes n-1} \otimes \mathbb{I}$, $\hat{U} = \sigma_x^{\otimes n-1} \otimes \mathbb{I}$, and $\check{U} = \sigma_z^{\otimes n-1} \otimes \mathbb{I}$ if n is even. It is easily verified that $\sigma_0^n = \bar{U}\bar{\sigma}_0^n\bar{U}^\dagger$, $\sigma_1^n = \bar{U}\bar{\sigma}_1^n\bar{U}^\dagger$, $\sigma_0^n = \hat{U}\hat{\sigma}_0^n\hat{U}^\dagger$, $\sigma_1^n = \hat{U}\hat{\sigma}_1^n\hat{U}^\dagger$, $\sigma_0^n = \check{U}\check{\sigma}_0^n\check{U}^\dagger$, and $\sigma_1^n = \check{U}\check{\sigma}_1^n\check{U}^\dagger$. We then get that

$$||\sigma_0^{n+2} - \sigma_1^{n+2}||_1 = ||\sigma_0^n - \sigma_1^n||_1$$

from which the claim follows. ∎

Surprisingly, if Bob does have some *a priori* knowledge about the outcome of the XOR, the problem becomes much harder for Bob. By expressing the states in the Bell basis and using Helstrom's result, it is easy to see that if Alice chooses $x \in \{0,1\}^2$ such that with probability q, $\mathrm{XOR}(x) = 0$, and with probability $(1-q)$, $\mathrm{XOR}(x) = 1$, Bob's probability of learning $\mathrm{XOR}(x)$ correctly is minimized for $q = 1/3$. In that case, Bob succeeds with probability at most $2/3$, which can be achieved by the trivial strategy of ignoring the state he received and always outputting 1. This is an explicit example where making a measurement does not aid in state discrimination. It has previously been noted by Hunter [25] that such cases can exist in mixed-state discrimination.

IV. Using Post-Measurement Information

We are now ready to advance to the core of our problem. Consider an instance of PI_0-STAR with a function $f: \mathcal{X} \to \mathcal{Y}$ and $m = |\mathcal{B}|$ bases, and some priors P_X and P_B on the sets $\mathcal{X}$ and $\mathcal{B}$. Since Bob may not store any quantum information, all his nontrivial actions are contained in the first measurement, which must equip him with possible outputs $o_i \in \mathcal{Y}$ for each basis $i = 1, \ldots, m$. In other words, his most general strategy is a positive operator-valued measure (POVM) with $|\mathcal{Y}|^m$ outcomes, each labeled by the strings $o_1, \ldots, o_m$ for $o_i \in \mathcal{Y}$ and $m = |\mathcal{B}|$. Once Alice has announced b, Bob outputs $\hat{Y} = o_b$. Here we first prove a general lower bound on the usefulness of post-measurement information that beats the guessing bound. Then, we analyze in detail the AND and the XOR function on n bits.

A. A Lower Bound for Balanced Functions

We first give a lower bound on Bob's success probability for any balanced function and any number of MUBs, by constructing an explicit measurement that achieves it. Without loss of generality, we assume in this section that $\mathcal{B} = [m]$, as otherwise we could consider a lexicographic ordering of $\mathcal{B}$.

Theorem IV.1: Let $f: \mathcal{X} \to \mathcal{Y}$ be a balanced function, and let P_X and P_B be the uniform distributions over $\mathcal{X}$ and $\mathcal{B}$, respectively. Let the set of unitaries $\{U_b | b \in \mathcal{B}\}$ give rise to $|\mathcal{B}|$ MUBs. Choose an encoding such that $\forall x, x' \in \mathcal{X} : \langle x|x'\rangle = \delta_{xx'}$. Then Bob succeeds at PI_0-$\mathrm{STAR}(f)$ with probability at least

$$p = p_{\mathrm{guess}} + \delta$$

with

$$\delta = \begin{cases} \frac{|\mathcal{Y}|-1}{|\mathcal{Y}|(|\mathcal{Y}|+3)}, & \text{if } m = 2 \\ \frac{4(|\mathcal{Y}|^2-1)}{3\mathcal{Y}(2+|\mathcal{Y}|(|\mathcal{Y}|+6))}, & \text{if } m = 3 \\ -\frac{2}{2|\mathcal{Y}|} + \frac{2(|\mathcal{Y}|+m-1)}{|\mathcal{Y}|^2+3|\mathcal{Y}|(m-1)+m^2-3m+2}, & \text{if } m \geq 4 \end{cases}$$

where p_{guess} is the probability that Bob can achieve by guessing the basis as given in Lemma II.5. In particular, we always have $p > p_{\mathrm{guess}}$.

Proof: Our proof works by constructing a square-root-type measurement that achieves the lower bound. As explained above, Bob's strategy for learning $f(x)$ is to perform a measurement with $|\mathcal{Y}|^m$ possible outcomes, labeled by the strings $o_1, \ldots, o_m$ for $o_i \in \mathcal{Y}$ and $m = |\mathcal{B}|$. Once Alice has announced b, Bob outputs $f(x) = o_b$.

Take the projector $P_{yb} = \sum_{x \in f^{-1}(y)} |\Phi_b^x\rangle\langle\Phi_b^x|$ and $\rho_{yb} = \frac{1}{k}P_{yb}$, where $k = |f^{-1}(y)| = |\mathcal{X}|/|\mathcal{Y}|$. Let $M_{o_1,\ldots,o_m}$ denote the measurement operator corresponding to outcome $o_1, \ldots, o_m$. Note that outcome $o_1, \ldots, o_m$ is the correct outcome for input state ρ_{yb} if and only if $o_b = y$. We can then write Bob's probability of success as

$$\frac{1}{m|\mathcal{Y}|} \sum_{o_1,\ldots,o_m \in \mathcal{Y}} \mathrm{Tr}\left(M_{o_1,\ldots,o_m}\left(\sum_{b \in B} \rho_{o_b b}\right)\right).$$

We will make use of the following measurement:

$$M_{o_1,\ldots,o_m} = S^{-\frac{1}{2}}\left(\sum_{b \in \mathcal{B}} P_{o_b b}\right)^3 S^{-\frac{1}{2}}$$

with

$$S = \sum_{o_1,\ldots,o_m \in \mathcal{Y}} \left(\sum_{b \in \mathcal{B}} P_{o_b b}\right)^3.$$

Clearly, we have $\sum_{o_1,\ldots,o_m\in\mathcal{Y}} M_{o_1,\ldots,o_m} = \mathbb{1}$ and $\forall o_1,\ldots,o_m \in \mathcal{Y} : M_{o_1,\ldots,o_m} \geq 0$ by construction and thus we indeed have a valid measurement. We first show that $S = c_m\mathbb{1}$.

$$\begin{aligned} S &= \sum_{o_1,\ldots,o_m\in\mathcal{Y}} \left(\sum_{b\in\mathcal{B}} P_{o_b b}\right)^3 \\ &= \sum_{o_1,\ldots,o_m\in\mathcal{Y}} \sum_{b,b',b''\in\mathcal{B}} P_{o_b b} P_{o_{b'} b'} P_{o_{b''} b''} \\ &= \Big[m|\mathcal{Y}|^{m-1} + 2m(m-1)|\mathcal{Y}|^{m-2} \\ &\quad + m(m-1)|\mathcal{Y}|^{m-2} \\ &\quad + m(m-1)(m-2)|\mathcal{Y}|^{m-3}\bar{\delta}_{2m}\Big]\mathbb{1} \end{aligned}$$

where we have used the assumption that for any b, $P_{o_b b}$ is a projector and $\sum_{x\in\mathcal{X}} |\Phi_b^x\rangle\langle\Phi_b^x| = \mathbb{1}$ which gives $\sum_{o_i\in\mathcal{Y}} P_{o_i b_i} = \sum_{o_i\in\mathcal{Y}} \sum_{x\in f^{-1}(y)} |\Phi_b^x\rangle\langle\Phi_b^x| = \mathbb{1}$. We can then write Bob's probability of success using this particular measurement as

$$\frac{1}{c_m k m |\mathcal{Y}|} \sum_{o_1,\ldots,o_m\in\mathcal{Y}} \mathrm{Tr}\left(\left(\sum_{b\in\mathcal{B}} P_{o_b b}\right)^4\right).$$

It remains to evaluate this expression. Using the circularity of the trace, we obtain

$$\begin{aligned} &\sum_{o_1,\ldots,o_m\in\mathcal{Y}} \mathrm{Tr}\left(\left(\sum_{b\in\mathcal{B}} P_{o_b b}\right)^4\right) \\ &\geq \Big[m|\mathcal{Y}|^{m-1} + 6m(m-1)|\mathcal{Y}|^{m-2} \\ &\quad + 6m(m-1)(m-2)|\mathcal{Y}|^{m-3}\bar{\delta}_{2m} \\ &\quad + m(m-1)(m-2)(m-3)|\mathcal{Y}|^{t(m-4)}\bar{\delta}_{2m}\bar{\delta}_{3m}\Big] \\ &\quad \times \mathrm{Tr}(\mathbb{1}) + m(m-1)|\mathcal{Y}|^{m-2}k \end{aligned}$$

where we have again used the assumption that for any b, $P_{o_b b}$ is a projector and $\sum_{x\in\mathcal{X}} |\Phi_b^x\rangle\langle\Phi_b^x| = \mathbb{1}$ with $\mathrm{Tr}(\mathbb{1}) = |\mathcal{X}|$. For the last term we have used the following: Note that $\mathrm{Tr}(P_{o_b b} P_{o_{b'} b'}) = k^2/|\mathcal{X}|$, because we assumed MUBs. Let $r = \mathrm{rank}(P_{o_b b} P_{o_{b'} b'})$. We can then bound $\mathrm{Tr}((P_{o_b b} P_{o_{b'} b'})^2) = \sum_i^r \lambda_i (P_{o_b b} P_{o_{b'} b'})^2 \geq k^4/(|\mathcal{X}|^2 r) \geq k^3/|\mathcal{X}|^2 = k/|\mathcal{Y}|^2$, where $\lambda_i(A)$ is the ith eigenvalue of a matrix A, by noting that $r \leq k$ since $\mathrm{rank}(P_{o_b b}) = \mathrm{rank}(P_{o'_b b'}) = k$. Putting things together we obtain

$$p \geq \frac{1}{c_m m}\left[G_m(1) + \left(6 + \frac{1}{|\mathcal{Y}|}\right) G_m(2) + 6G_m(3) + G_m(4)\right]$$

where $m = |\mathcal{B}|$, $c_m = G_m(1) + 3G_m(2) + G_m(3)$ and function $G_m : \mathbb{N} \to \mathbb{N}$ defined as

$$G_m(i) = \frac{m!}{(m-i)!} |\mathcal{Y}|^{m-i} \prod_{j=2}^{i-1} \bar{\delta}_{mj}.$$

This expression can be simplified to obtain the claimed result. ∎

Note that we have only used the assumption that Alice uses MUBs in the very last step to say that $\mathrm{Tr}(P_{o_b b} P_{o_{b'} b'}) = k^2/|\mathcal{X}|$. One could generalize our argument to other cases by evaluating $\mathrm{Tr}(P_{o_b b} P_{o_{b'} b'})$ approximately. In the special case $m = |\mathcal{Y}| = 2$ (i.e., binary function, with two bases) we obtain the following.

Corollary IV.2: Let $f: \{0,1\}^n \to \{0,1\}$ be a balanced function and let $P_X(x) = 2^{-n}$ for all $x \in \{0,1\}^n$. Let $\mathcal{B} = \{0,1\}$ with $U_0 = \mathbb{1}^{\otimes n}$, $U_1 = H^{\otimes n}$, and $P_B(0) = P_B(1) = 1/2$. Then Bob succeeds at PI_0-$\mathrm{STAR}(f)$ with probability $p \geq 0.85$. □

Observe that this almost attains the upper bound of ≈ 0.853 of Lemma III.1 in the case of no post-measurement information. Below (in Section VI) we will show that indeed this bound can always be achieved when post-measurement information is available.

It is perhaps interesting to note that our general bound depends only on the number of function values $|\mathcal{Y}|$ and the number of bases m. The number of function inputs $|\mathcal{X}|$ itself does not play a direct role.

B. Optimal Bounds for the AND and XOR Functions

We now show that for some specific functions, the probability of success can even be much larger. We hereby concentrate on the case where Alice uses two or three MUBs to encode her input. Our proofs thereby lead to explicit measurements. In the following, we again assume that Bob has no *a priori* knowledge of the function value.

1) AND *Function:*

Theorem IV.3: Let $P_X(x) = \frac{1}{2(2^n-1)}$ for all $x \in \{0,1\}^n \setminus \{1\ldots1\}$ and $P_X(1\ldots1) = \frac{1}{2}$. Let $\mathcal{B} = \{+,\times\}$ with $U_+ = \mathbb{1}^{\otimes n}$, $U_\times = H^{\otimes n}$, and $P_B(+) = P_B(\times) = 1/2$. Then Bob succeeds at PI_0-$\mathrm{STAR}(\mathrm{AND})$ with probability at most

$$p = \frac{1}{2}\left[2 + \frac{1}{2^n + 2^{n/2} - 2} - \frac{1}{2^n - 1}\right]. \tag{8}$$

There exists a strategy for Bob that achieves p.

Proof: To learn the value of $\mathrm{AND}(x)$, Bob uses the same strategy as in the previous section. He performs a measurement with four possible outcomes, labeled by the strings $o_+, o_\times$ with $o_+, o_\times \in \{0,1\}$. Once Alice has announced her basis choice $b \in \{+,\times\}$, Bob outputs $\mathrm{AND}(x) = o_b$. Note that w.l.o.g. we can assume that Bob's measurement has only four outcomes, i.e., Bob only stores 2 bits of classical information because he will only condition his answer on the value of b later on.

Following the approach in the last section, we can write Bob's optimal probability of success as a semidefinite program

$$\begin{aligned} \text{maximize} \quad & \frac{1}{4} \sum_{o_+,o_\times\in\{0,1\}} \mathrm{Tr}[b_{o_+ o_\times} M_{o_+ o_\times}] \\ \text{subject to} \quad & \forall o_+, o_\times \in \{0,1\} : M_{o_+ o_\times} \geq 0, \\ & \sum_{o_+,o_\times\in\{0,1\}} M_{o_+ o_\times} = \mathbb{1} \end{aligned}$$

where

$$\begin{aligned} b_{00} &= \rho_{0+} + \rho_{0\times}, b_{01} = \rho_{0+} + \rho_{1\times} \\ b_{10} &= \rho_{1+} + \rho_{0\times}, b_{11} = \rho_{1+} + \rho_{1\times} \end{aligned}$$

with $\forall y \in \{0,1\}$

$$b \in \{+,\times\} : \rho_{yb} = \frac{1}{2} \sum_{x\in\mathrm{AND}^{-1}(y)} U_b |x\rangle\langle x| U_B^\dagger.$$

Consider $\mathcal{H}_2$, the two-dimensional Hilbert space spanned by $|c_1\rangle \stackrel{\text{def}}{=} |1\rangle^{\otimes n}$ and $|h_1\rangle \stackrel{\text{def}}{=} |1_H\rangle^{\otimes n}$. Let $|c_0\rangle \in \mathcal{H}_2$ and $|h_0\rangle \in \mathcal{H}_2$ be the state vectors orthogonal to $|c_1\rangle$ and $|h_1\rangle$, respectively. They can be expressed as

$$|c_o\rangle = \frac{(-1)^{n+1}|c_1\rangle + 2^{n/2}|h_1\rangle}{\sqrt{2^n - 1}}$$
$$|h_o\rangle = \frac{2^{n/2}|c_1\rangle + (-1)^{n+1}|h_1\rangle}{\sqrt{2^n - 1}}.$$

Then $\Pi_{\|} = |c_0\rangle\langle c_0| + |c_1\rangle\langle c_1| = |h_0\rangle\langle h_0| + |h_1\rangle\langle h_1|$ is a projector onto $\mathcal{H}_2$. Let $\Pi_{\perp}$ be a projector onto the orthogonal complement of $\mathcal{H}_2$. Note that the $b_{o_+o_\times}$ are all composed of two blocks, one supported on $\mathcal{H}_2$ and the other on its orthogonal complement. We can thus write

$$\begin{aligned} b_{00} &= \frac{2\Pi_{\perp}}{2^n - 1} + \frac{|c_0\rangle\langle c_0| + |h_0\rangle\langle h_0|}{2^n - 1} \\ b_{01} &= \frac{\Pi_{\perp}}{2^n - 1} + \left[\frac{|c_0\rangle\langle c_0|}{2^n - 1} + |h_1\rangle\langle h_1|\right] \\ b_{10} &= \frac{\Pi_{\perp}}{2^n - 1} + \left[\frac{|h_0\rangle\langle h_0|}{2^n - 1} + |c_1\rangle\langle c_1|\right] \\ b_{11} &= 0 + |c_1\rangle\langle c_1| + |h_1\rangle\langle h_1|. \end{aligned} \tag{9}$$

We give an explicit measurement that achieves p and then show that it is optimal. The full derivation of this measurement can be found in the Appendix. Take

$$M_{00} = \Pi_{\perp}$$
$$M_{o_+o_\times} = \lambda_{o_+o_\times}|\psi_{o_+o_\times}\rangle\langle\psi_{o_+o_\times}|$$

with $\lambda_{01} = \lambda_{10} = (1+\eta)^{-1}$, where

$$\eta = \left|\frac{1 - 2\beta^2 + (-1)^{n+1}2\beta\sqrt{1-\beta^2}\sqrt{2^n - 1}}{2^{n/2}}\right|$$
$$|\psi_{01}\rangle = \alpha|c_0\rangle + \beta|c_1\rangle$$
$$|\psi_{10}\rangle = \alpha|h_0\rangle + \beta|h_1\rangle$$

with α and β real and satisfying $\alpha^2 + \beta^2 = 1$. We also set $M_{11} = \mathbb{I} - M_{00} - M_{01} - M_{10}$. We take

$$\beta = (-1)^n \frac{1}{\sqrt{2^{2n} + 2^{\frac{3}{2}n+1} - 2^{\frac{n}{2}+1}}}.$$

Putting it all together, we thus calculate Bob's probability of success

$$p = \frac{1}{2}\left[2 + \frac{1}{2^n + 2^{n/2} - 2} - \frac{1}{2^n - 1}\right].$$

We now show that this is in fact the optimal measurement for Bob. For this we will consider the dual of our semidefinite program above

$$\begin{aligned} &\text{minimize} \quad \operatorname{Tr}(Q) \\ &\text{subject to} \quad \forall o_+, o_\times \in \{0,1\} : Q \geq \frac{b_{o_+o_\times}}{4}. \end{aligned}$$

Our goal is now to find a Q such that $p = \operatorname{Tr}(Q)$ and Q is dual feasible. We can then conclude from the duality of SDP that p is optimal. Consider

$$\begin{aligned} Q = &\frac{\Pi_{\perp}}{2(2^n - 1)} \\ &+ \frac{1}{4}\left(\frac{2 - 2^{1+n/2} + 2^{3n/2}}{2 - 3\cdot 2^{n/2} + 2^{3n/2}}\right)(|c_1\rangle\langle c_1| + |h_1\rangle\langle h_1|) \\ &- (-1)^n \frac{1}{4(2^{1-\frac{n}{2}} + 2^n - 3)}(|c_1\rangle\langle h_1| + |c_1\rangle\langle h_1|). \end{aligned}$$

Now we only need to show that the Q above satisfies the constraints, i.e., $\forall o_+, o_\times \in \{0,1\} : Q \geq b_{o_+o_\times}/4$. Let $Q_{\perp} = \Pi_{\perp} Q \Pi_{\perp}$ and $Q_{\|} = \Pi_{\|} Q \Pi_{\|}$. By taking a look at (9) one can easily see that $Q_{\perp} \geq \frac{\Pi_{\perp} b_{o_+o_\times} \Pi_{\perp}}{4}$, so that it is only left to show that

$$Q_{\|} \geq \frac{\Pi_{\|} b_{o_+o_\times} \Pi_{\|}}{4}, \quad \text{for } o_+o_\times \in \{0,1\}, o_+o_\times \neq 00.$$

These are 2×2 matrices and this can be done straightforwardly. We thus have $\operatorname{Tr}(Q) = p$ and the result follows from the duality of semidefinite programming. ∎

It also follows that if Bob just wants to learn the value of a single bit, he can do no better than what he could achieve without waiting for Alice's announcement of the basis b.

Corollary IV.4: Let $x \in \{0,1\}$, $P_X(x) = \frac{1}{2}$, and $f(x) = x$. Let $\mathcal{B} = \{+, \times\}$ with $U_+ = \mathbb{I}$ and $U_\times = H$. Then Bob succeeds at $\text{PI}_0\text{-STAR}(f)$ with probability at most

$$p = \frac{1}{2} + \frac{1}{2\sqrt{2}}.$$

There exists a strategy for Bob that achieves p. □

The AND function provides an intuitive example of how Bob can compute the value of a function perfectly by storing just a single qubit. Consider the measurement with elements $\{\Pi_{\|}, \Pi_{\perp}\}$ from the previous section. It is easy to see that the outcome $\perp$ has zero probability if $\text{AND}(x) = 1$. Thus, if Bob obtains that outcome he can immediately conclude that $\text{AND}(x) = 0$. If Bob obtains outcome $\|$ then the post-measurement states live in a two-dimensional Hilbert space $(\mathcal{H}_2)$, and can therefore be stored in a single qubit. Thus, by keeping the remaining state we can calculate the AND perfectly once the basis is announced. Our proof in Section VI, which shows that in fact *all* Boolean functions can be computed perfectly if Bob can store only a single qubit, makes use of a very similar effect to the one we observed here explicitly.

2) XOR *Function:* We now examine the XOR function. This will be useful in order to gain some insight into the usefulness of post-measurement information later. For strings of even length, there exists a simple strategy for Bob even when three MUBs are used.

Theorem IV.5: Let $n \in \mathbb{N}$ be even, and let $P_X(x) = \frac{1}{2^n}$ for all $x \in \{0,1\}^n$. Let $\mathcal{B} = \{+, \times, \odot\}$ with $U_+ = \mathbb{I}^{\otimes n}$, $U_\times = H^{\otimes n}$, and $U_\odot = K^{\otimes n}$, where $K = (\mathbb{I} + i\sigma_x)/\sqrt{2}$. Then

there is a strategy where Bob succeeds at PI_0-$\mathrm{STAR}(\mathrm{XOR})$ with probability $p = 1$.

Proof: We first construct Bob's measurement for the first two qubits, which will allow him to learn $x_1 \oplus x_2$ with probability 1. Note that the 12 possible states that Alice sends can be expressed in the Bell basis as follows:

$$\begin{aligned}
|00\rangle &= \frac{1}{\sqrt{2}}(|\Phi^+\rangle + |\Phi^-\rangle)\\
|01\rangle &= \frac{1}{\sqrt{2}}(|\Psi^+\rangle + |\Psi^-\rangle)\\
|10\rangle &= \frac{1}{\sqrt{2}}(|\Psi^+\rangle - |\Psi^-\rangle)\\
|11\rangle &= \frac{1}{\sqrt{2}}(|\Phi^+\rangle - |\Phi^-\rangle)\\
H^{\otimes 2}|00\rangle &= \frac{1}{\sqrt{2}}(|\Phi^+\rangle + |\Psi^+\rangle)\\
H^{\otimes 2}|01\rangle &= \frac{1}{\sqrt{2}}(|\Phi^-\rangle + |\Psi^-\rangle)\\
H^{\otimes 2}|10\rangle &= \frac{1}{\sqrt{2}}(|\Phi^-\rangle - |\Psi^-\rangle)\\
H^{\otimes 2}|11\rangle &= \frac{1}{\sqrt{2}}(|\Phi^+\rangle - |\Psi^+\rangle)\\
K^{\otimes 2}|00\rangle &= \frac{1}{\sqrt{2}}(|\Phi^-\rangle + i|\Psi^+\rangle)\\
K^{\otimes 2}|01\rangle &= \frac{1}{\sqrt{2}}(i|\Phi^+\rangle + |\Psi^-\rangle)\\
K^{\otimes 2}|10\rangle &= \frac{1}{\sqrt{2}}(i|\Phi^+\rangle - |\Psi^-\rangle)\\
K^{\otimes 2}|11\rangle &= -\frac{1}{\sqrt{2}}(|\Phi^-\rangle - i|\Psi^+\rangle).
\end{aligned}$$

Bob now simply measures in the Bell basis and records his outcome. If Alice now announces that she used the computational basis, Bob concludes that $x_1 \oplus x_2 = 0$ if the outcome is one of $|\Phi^{\pm}\rangle$ and $x_1 \oplus x_2 = 1$ otherwise. If Alice announces she used the Hadamard basis, Bob concludes that $x_1 \oplus x_2 = 0$ if the outcome was one of $\{|\Phi^+\rangle, |\Psi^+\rangle\}$ and $x_1 \oplus x_2 = 1$ otherwise. Finally, if Alice announces that she used the $\odot$ basis, Bob concludes that $x_1 \oplus x_2 = 0$ if the outcome was one of $\{|\Phi^-\rangle, |\Psi^+\rangle\}$ and $x_1 \oplus x_2 = 1$ otherwise. Bob can thus learn the XOR of two bits with probability 1. To learn the XOR of the entire string, Bob applies this strategy to each two bits individually and then computes the XOR of all answers. ∎

Analogously to the proof of Theorem IV.5, we obtain the following.

Corollary IV.6: Let $n \in \mathbb{N}$ be even, and let $P_X(x) = \frac{1}{2^n}$ for all $x \in \{0,1\}^n$. Let $\mathcal{B} = \{+, \times\}$ with $U_+ = \mathbb{I}^{\otimes n}$ and $U_\times = H^{\otimes n}$. Then there is a strategy where Bob succeeds at PI_0-$\mathrm{STAR}(\mathrm{XOR})$ with probability $p = 1$. □

Interestingly, there is no equivalent strategy for Bob if n is odd. In fact, as we will show in the next section, in this case the post-measurement information gives no advantage to Bob at all.

Theorem IV.7: Let $n \in \mathbb{N}$ be odd, and let $P_X(x) = \frac{1}{2^n}$ for all $x \in \{0,1\}^n$. Let $\mathcal{B} = \{+, \times\}$ with $U_+ = \mathbb{I}^{\otimes n}$, $U_\times = H^{\otimes n}$, and $P_B(+) = P_B(\times) = 1/2$. Then Bob succeeds at PI_0-$\mathrm{STAR}(\mathrm{XOR})$ with probability at most

$$p = \frac{1}{2}\left(1 + \frac{1}{\sqrt{2}}\right).$$

There exists a strategy for Bob that achieves p.

Proof: Similar to the proof of the AND function, we can write Bob's optimal probability of success as the following semidefinite program in terms of the length of the input string n:

$$\begin{aligned}
\text{maximize} \quad & \frac{1}{4}\sum_{o_+,o_\times \in \{0,1\}} \mathrm{Tr}[b^n_{o_+o_\times} M_{o_+o_\times}]\\
\text{subject to} \quad & \forall o_+, o_\times \in \{0,1\} : M_{o_+o_\times} \geq 0,\\
& \sum_{o_+,o_\times \in \{0,1\}} M_{o_+o_\times} = \mathbb{I}
\end{aligned}$$

where

$$b^n_{o_+o_\times} = \rho^n_{o_+ +} + \rho^n_{o_\times \times}$$

and $\rho^n_{o_b b} = \frac{1}{2^{n-1}}\sum_{x \in \{0,1\}^n, x \in \mathrm{XOR}^{-1}(o_b)} U_b|x\rangle\langle x|U_b^\dagger$. The dual can be written as

$$\begin{aligned}
\text{minimize} \quad & \frac{1}{4}\mathrm{Tr}(Q^n)\\
\text{subject to} \quad & \forall o_+, o_\times \in \{0,1\} : Q^n \geq b^n_{o_+o_\times}.
\end{aligned}$$

Our proof is now by induction on n. For $n = 1$, let $Q^1 = 2p\mathbb{I}$. It is easy to verify that $\forall o_+, o_\times \in \{0,1\} : Q^1 \geq b^1_{o_+o_\times}$ and thus Q^1 is a feasible solution of the dual program.

We now show that for $n+2$, $Q^{n+2} = Q^n \otimes \frac{1}{4}\mathbb{I}$ is a feasible solution to the dual for $n+2$, where Q^n is a solution for the dual for n. Note that the XOR of all bits in the string can be expressed as the XOR of the first $n-2$ bits XORed with the XOR of the last two. Recall (7). Now note that we can write

$$\begin{aligned}
\rho^2_{0+} &= \frac{1}{2}(|00\rangle\langle 00| + |11\rangle\langle 11|)\\
&= \frac{1}{2}(|\Phi^+\rangle\langle\Phi^+| + |\Phi^-\rangle\langle\Phi^-|)\\
\rho^2_{1+} &= \frac{1}{2}(|01\rangle\langle 01| + |10\rangle\langle 10|)\\
&= \frac{1}{2}(|\Psi^+\rangle\langle\Psi^+| + |\Psi^-\rangle\langle\Psi^-|).
\end{aligned}$$

It is easy to see that $\rho^2_{0\times} = H\rho^2_{0+}H = \frac{1}{2}(|\Phi^+\rangle\langle\Phi^+| + |\Psi^+\rangle\langle\Psi^+|)$ and $\rho^2_{1\times} = H\rho^2_{1+}H = \frac{1}{2}(|\Phi^-\rangle\langle\Phi^-| + |\Psi^-\rangle\langle\Psi^-|)$. By substituting from the above equation we then obtain

$$\begin{aligned}
b^{n+2}_{00} &= \rho^{n+2}_{0+} + \rho^{n+2}_{0\times}\\
&= \frac{1}{4}\left((\rho^n_{0+} + \rho^n_{0\times}) \otimes |\Phi^+\rangle\langle\Phi^+| + (\rho^n_{0+} + \rho^n_{1\times})\right.\\
&\quad \otimes |\Phi^-\rangle\langle\Phi^-| (\rho^n_{1+} + \rho^n_{0\times}) \otimes |\Psi^+\rangle\langle\Psi^+|\\
&\quad \left. + (\rho^n_{1+} + \rho^n_{1\times}) \otimes |\Psi^-\rangle\langle\Psi^-|\right)\\
&\leq \frac{1}{4}Q^n \otimes \mathbb{I}
\end{aligned}$$

where we have used the fact that Q^n is a feasible solution for the dual for n and that $|\Phi^+\rangle\langle\Phi^+| + |\Phi^-\rangle\langle\Phi^-| + |\Psi^+\rangle\langle\Psi^+| + |\Psi^-\rangle\langle\Psi^-| = \mathbb{I}$. The argument for b^{n+2}_{01}, b^{n+2}_{10}, and b^{n+2}_{11} is analogous. Thus, Q^{n+2} satisfies all constraints.

Putting things together, we have for odd n that $\mathrm{Tr}(Q^{n+2}) = \mathrm{Tr}(Q^n) = \mathrm{Tr}(Q^1)$ and since the dual is a minimization problem

we know that $p \leq \frac{1}{4}\mathrm{Tr}(Q^1) = c$ as claimed. Clearly, there exists a strategy for Bob that achieves $p = c$. He can compute the XOR of the first $n-1$ bits perfectly, as shown in Theorem IV.6. By Corollary IV.4, he can learn the value of the remaining nth bit with probability $p = c$. ∎

We obtain a similar bound for three bases.

Theorem IV.8: Let $n \in \mathbb{N}$ be odd, and let $P_X(x) = \frac{1}{2^n}$ for all $x \in \{0,1\}^n$. Let $\mathcal{B} = \{+, \times, \odot\}$ with $U_+ = \mathbb{1}^{\otimes n}$, $U_\times = H^{\otimes n}$, and $U_\odot = K^{\otimes n}$, where $K = (\mathbb{1} + i\sigma_x)/\sqrt{2}$, with $P_B(+) = P_B(\times) = P_B(\odot) = 1/3$. Then Bob succeeds at PI_0-STAR(XOR) with probability at most

$$p = \frac{1}{2}\left(1 + \frac{1}{\sqrt{3}}\right).$$

There exists a strategy for Bob that achieves p.

Proof: The proof follows the same lines as Theorem IV.7. Bob's optimal probability of success is

$$\begin{aligned} \text{maximize} \quad & \frac{1}{6}\sum_{o_+,o_\times,o_\odot \in \{0,1\}} \mathrm{Tr}[b^n_{o_+o_\times o_\odot} M_{o_+o_\times o_\odot}] \\ \text{subject to} \quad & \forall o_+, o_\times, o_\odot \in \{0,1\} \in \{0,1\} : M_{o_+o_\times o_\odot} \geq 0, \\ & \sum_{o_+,o_\times,o_\odot \in \{0,1\}} M_{o_+o_\times o_\odot} = \mathbb{1} \end{aligned}$$

where

$$b^n_{o_+o_\times o_\odot} = \sum_{b \in \mathcal{B}} \rho_{o_b b},$$

and

$$\rho_{o_b b} = \frac{1}{2^{n-1}} \sum_{x \in \mathrm{XOR}(o_b)} U_b |x\rangle\langle x| U_b^\dagger.$$

The dual can be written as

$$\begin{aligned} \text{minimize} \quad & \frac{1}{6}\mathrm{Tr}(Q^n) \\ \text{subject to} \quad & \forall o_+, o_\times, o_\odot \in \{0,1\} : Q^n \geq b^n_{o_+o_\times o_\odot}. \end{aligned}$$

Again, the proof continues by induction on n. For $n = 1$, let $Q^1 = 3p\mathbb{1}$. It is easy to verify that $\forall o_+, o_\times, o_\odot \in \{0,1\} : Q^1 \geq b^1_{o_+o_\times o_\odot}$ and, thus, Q^1 is a feasible solution of the dual program. The rest of the proof is done exactly in the same way as in Theorem IV.7 using that

$$\begin{aligned} \rho^2_{0\odot} &= \frac{1}{2}(|\Phi^-\rangle\langle\Phi^-| + |\Psi^+\rangle\langle\Psi^+|) \\ \rho^2_{1\odot} &= \frac{1}{2}(|\Psi^-\rangle\langle\Psi^-| + |\Phi^+\rangle\langle\Phi^+|). \end{aligned}$$ ∎

V. Quantum Memory Resources for Perfect Prediction: An Algebraic Framework

So far, we had assumed that Bob is not allowed to store any qubits and can only use the additional post-measurement information to improve his guess. Now, we investigate the case where he has a certain amount of quantum memory at his disposal. In particular, we present a general algebraic approach to determine the minimum dimension 2^q of quantum memory needed to succeed with probability 1 at an instance of PI_q-STAR($\mathcal{E}$) for any ensemble $\mathcal{E} = \{p_{yb}, \rho_{yb}\}$ as long as the individual states for different values of y are mutually orthogonal for a fixed b, i.e., $\forall y \neq z \in \mathcal{Y} : \mathrm{Tr}(\rho_{yb}\rho_{zb}) = 0$. We are looking for an instrument (w.l.o.g. maximally refined, since from the rank bound it is clear that randomization or using impure completely positive maps in the instrument does not result in any advantage) consisting of a family of pure completely positive maps $\rho \mapsto A\rho A^\dagger$, adding up to a trace preserving map, such that $\mathrm{rank}\, A \leq 2^q$. This takes care of the memory bound. The fact that after the announcement of b the remaining state $A\rho_{yb}A^\dagger$ gives full information about y is expressed by demanding orthogonality of the different post-measurement states

$$\forall b \in \mathcal{B}, \forall y \neq z \in \mathcal{Y}, \quad A\rho_{yb}A^\dagger A\rho_{zb}A^\dagger = 0. \tag{10}$$

Note that here we explicitly allow the possibility that, say, $A\rho_{zb}A^\dagger = 0$: this means that if Bob obtains outcome A and later learns b, he can exclude the output value z. What (10) also implies is that for all states $|\psi\rangle$ and $|\varphi\rangle$ in the support of ρ_{yb} and ρ_{zb}, respectively, on has $A|\psi\rangle\langle\psi|A^\dagger A|\varphi\rangle\langle\varphi|A^\dagger = 0$, hence, introducing the support projectors P_{yb} of the ρ_{yb}, we can reformulate (10) as $\forall b \in \mathcal{B}, \forall y \neq z \in \mathcal{Y}$

$$AP_{yb}A^\dagger AP_{zb}A^\dagger = 0$$

which can equivalently be expressed as $\forall b \in \mathcal{B}, \forall y \neq z \in \mathcal{Y}$

$$\mathrm{Tr}\left(A^\dagger AP_{yb}A^\dagger AP_{zb}\right) = 0. \tag{11}$$

As expected, we see that only the POVM operators $M = A^\dagger A$ of the instrument play a role in this condition. Our conditions can therefore also be written as $MP_{yb}MP_{zb} = 0$. From this condition, we now derive the following lemma.

Lemma V.1: Bob, using an instrument with POVM operators $\{M_i\}$, succeeds at PI_q-STAR with probability 1, if and only if

1) for all i, $\mathrm{rank}\, M_i \leq 2^q$;
2) for all $y \in \mathcal{Y}$ and $b \in \mathcal{B}$, $[M, P_{yb}] = 0$, where P_{yb} is the projection on the support of ρ_{yb}.

Proof: We first show that these two conditions are necessary. Note that only the commutation has to be proved: let M be a Kraus element from an instrument succeeding with probability 1. Then, for any y, b, we have by (11) that

$$\mathrm{Tr}\left(MP_{yb}M(\mathbb{1} - P_{yb})\right) = 0$$

hence

$$\mathrm{Tr}\left(MP_{yb}MP_{yb}\right) = \mathrm{Tr}\left(MP_{yb}M\right).$$

Thus, by the positivity of the trace on positive operators, the cyclicity of the trace, and $P^2_{yb} = P_{yb}$ we have that

$$\begin{aligned} 0 &\leq \mathrm{Tr}\big([M, P_{yb}]^\dagger [M, P_{yb}]\big) \\ &= \mathrm{Tr}\big(-(MP_{yb} - P_{yb}M)^2\big) \\ &= \mathrm{Tr}\big(-MP_{yb}MP_{yb} - P_{yb}MP_{yb}M + P_{yb}M^2P_{yb} + MP^2_{yb}M\big) \\ &= 0. \end{aligned}$$

But that means that the commutator $[M, P_{yb}]$ has to be 0.

Sufficiency is easy: since the measurement operators commute with the states' support projectors P_{yb} (assuming for the

moment that they are the signals, not the ρ_{yb}), and these are orthogonal to each other for fixed b, the post-measurement states of these projectors $\propto \sqrt{M}P_{yb}\sqrt{M}$ will also be mutually orthogonal for fixed b. Thus, if Bob learns b, he can perform a measurement to distinguish the different values of y perfectly. The post-measurement states are clearly supported on the support of M, which can be stored in q qubits. Since Bob's strategy succeeds with probability 1, it will succeed with probability 1 for any states supported in the range of the P_{yb}. ∎

It should be pointed out that the operators M of the instrument need not commute with the originally given states ρ_{yb}. Nevertheless, the measurement preserves the orthogonality of ρ_{yb} and ρ_{zb} with $y \neq z$ for fixed b, i.e., $\mathrm{Tr}(\rho_{yb}\rho_{zb}) = 0$. Now that we know that the POVM operators of the instrument have to commute with all the states' support projectors P_{yb}, we can invoke some well-developed algebraic machinery to find the optimal such instrument.

Namely, the M have to come from the commutant of the operators P_{yb} [12]. These themselves generate a $*$-subalgebra $\mathcal{O}$ of the full operator algebra $\mathcal{B}(\mathcal{H})$ of the underlying Hilbert space $\mathcal{H}$, and the structure of such algebras and their commutants in finite dimension is well understood [33, Sec. I.II]: the Hilbert space $\mathcal{H}$ has a decomposition (i.e., there is an isomorphism which we write as an equality)

$$\mathcal{H} = \bigoplus_j \mathcal{J}_j \otimes \mathcal{K}_j \tag{12}$$

into a direct sum of tensor products, such that the $*$-algebra $\mathcal{O}$ and its commutant algebra $\mathcal{O}' = \{M : \forall P \in \mathcal{O}\ [P, M] = 0\}$ can be written

$$\mathcal{O} = \bigoplus_j \mathcal{B}(\mathcal{J}_j) \otimes \mathbb{1}_{\mathcal{K}_j} \tag{13}$$

$$\mathcal{O}' = \bigoplus_j \mathbb{1}_{\mathcal{J}_j} \otimes \mathcal{B}(\mathcal{K}_j). \tag{14}$$

Koashi and Imoto [28], in the context of finding the quantum operations which leave a set of states invariant, have described an algorithm to find the commutant $\mathcal{O}'$, and more precisely the Hilbert space decomposition (12), of the states $P_{yb}/\mathrm{Tr}P_{yb}$. They show that for this decomposition, there exist states $\sigma_{j|i}$ on $\mathcal{J}_j$, a conditional probability distribution $\{q_{j|i}\}$, and states ω_j on $\mathcal{K}_j$ which are independent of i, such that we can write them as

$$\forall i, \quad \sigma_i = \bigoplus_j q_{j|i}\sigma_{j|i} \otimes \omega_j,$$

Now, looking at (14), we see that the smallest rank operators $M \in \mathcal{O}'$ are of the form $\mathbb{1}_{\mathcal{J}_j} \otimes |\psi\rangle\langle\psi|$ for some j and $|\psi\rangle \in \mathcal{K}_j$, and that they are all admissible. Since we need a family of operators M that are closed to a POVM and thus all j have to occur, the minimal quantum memory requirement is

$$\min 2^q = \max_j \dim \mathcal{J}_j. \tag{15}$$

The strategy Bob has to follow is this: For each j, pick a basis $\{|e_{k|j}\rangle\}$ of the spaces $\mathcal{K}_j$ and measure the POVM $\{\mathbb{1}_{\mathcal{J}_j} \otimes |e_{k|j}\rangle\langle e_{k|j}|\}$, corresponding to the decomposition

$$\mathcal{H} = \bigoplus_{jk} \mathcal{J}_j \otimes |e_{k|j}\rangle$$

which commutes with the P_{yb}. For each outcome, he can store the post-measurement state in q qubits (as in (15)), preserving the orthogonality of the states for different y but fixed b. Once he learns b he can thus obtain y with certainty.

Of course, carrying out the Koashi–Imoto algorithm may not be a straightforward task in a given situation. Nevertheless, one can understand the two examples we will present in the following section as special cases of this general method.

VI. Using Post-Measurement Information and Quantum Memory

We now take a look at two specific cases. First, we show that in fact *all* Boolean functions with two bases (mutually unbiased or not) can be computed perfectly when Bob is allowed to store just a single qubit. Second, however, we show that there exist three bases such that for *any balanced* function, Bob must store *all* qubits to compute the function perfectly. We also give a recipe how to construct such bases.

A. Using Two Bases

For two bases, Bob needs to store only a single qubit to compute any Boolean function perfectly. As outlined in Section V, we need to show that there exists a measurement with the following properties: First, the posterior states of states corresponding to strings x such that $f(x) = 0$ are orthogonal to the posterior states of states corresponding to strings y such that $f(y) = 1$. Indeed, if this is true and we keep the posterior state, then after the basis is announced we can distinguish perfectly between both types of states. Second, of course, we need that the posterior states are supported in subspaces of dimension at most 2. The following lemma is the main ingredient in our proof. The same statement has been proven by Masanes [30] in a different context.

Lemma VI.1: Let $f\colon \{0,1\}^n \to \{0,1\}$ and

$$P_{0b} = \sum_{x \in f^{-1}(0)} U_b|x\rangle\langle x|U_b^\dagger$$

where $U_0 = \mathbb{1}$ and $U_1 = U$, then there exists a direct sum decomposition of the Hilbert space

$$\mathcal{H} = \bigoplus_{i=1}^m \mathcal{H}_i, \quad \text{with } \dim\mathcal{H}_i \leq 2$$

such that P_{00} and P_{01} can be expressed as

$$P_{00} = \sum_{i=1}^m \Pi_i P_{00} \Pi_i$$

$$P_{01} = \sum_{i=1}^m \Pi_i P_{01} \Pi_i$$

where Π_i is the orthogonal projector onto $\mathcal{H}_i$.

Proof: There exists a basis so that P_{00} and P_{01} can be written as

$$P_{00} = \begin{bmatrix} \mathbb{1}_{n_0} & 0_{n_0 \times n_1} \\ 0_{n_1 \times n_0} & 0_{n_1 \times n_1} \end{bmatrix}$$

$$P_{01} = \begin{bmatrix} A^{00}_{n_0 \times n_0} & A^{01}_{n_0 \times n_1} \\ (A^{01})^\dagger_{n_1 \times n_0} & A^{11}_{n_1 \times n_1} \end{bmatrix}$$

where $n_y = |f^{-1}(y)|$ is the number of strings x such that $f(x) = y$, and we have specified the dimensions of the matrix blocks for clarity. In what follows these dimensions will be omitted. We assume w.l.o.g. that $n_0 \leq n_1$. It is easy to check that, since P_{01} is a projector, it must satisfy

$$A^{00}(\mathbb{I}_{n_0} - A^{00}) = A^{01}A^{01\dagger}$$
$$A^{11}(\mathbb{I}_{n_1} - A^{11}) = A^{01\dagger}A^{01}. \tag{16}$$

Consider a unitary of the following form:

$$V = \begin{bmatrix} V_0 & 0 \\ 0 & V_1 \end{bmatrix}$$

where V_0 and V_1 are $n_0 \times n_0$ and $n_1 \times n_1$ unitaries, respectively. Under such a unitary, P_{00} and P_{01} are transformed to

$$VP_{00}V^\dagger = P_{00} \tag{17}$$
$$VP_{01}V^\dagger = \begin{bmatrix} V_0A^{00}V_0^\dagger & V_0A^{01}V_1^\dagger \\ (V_0A^{01}V_1^\dagger)^\dagger & V_1A^{11}V_1^\dagger \end{bmatrix}. \tag{18}$$

We now choose V_0 and V_1 from the singular value decomposition (SVD, [24, Theorem 7.3.5]) of $A^{01} = V_0^\dagger D V_1$ which gives

$$D = V_0A^{01}V_1^\dagger = \sum_{k=1}^{n_0} d_k|u_k\rangle\langle v_k|$$

where $d_k \geq 0$, $\langle u_k u_l\rangle = \langle v_k|v_l\rangle = \delta_{kl}$. Since $(A^{01})^\dagger A^{01}$ and $A^{01}(A^{01})^\dagger$ are supported in orthogonal subspaces, it also holds that $\forall k, l : \langle u_k|v_l\rangle = 0$. Equations (16), (17), and (18) now give us

$$V_0A^{00}V_0^\dagger(\mathbb{I}_{n_0} - V_0A^{00}V_0^\dagger) = \sum_{k=1}^{n_0} d_k^2|u_k\rangle\langle u_k|$$
$$V_1A^{11}V_1^\dagger(\mathbb{I}_{n_1} - V_1A^{11}V_1^\dagger) = \sum_{k=1}^{n_0} d_k^2|v_k\rangle\langle v_k|.$$

Suppose for the time being that all the d_k are different. Since they are all nonnegative then all the d_k^2 will also be different and it must hold that

$$V_0A^{00}V_0^\dagger = \sum_{k=1}^{n_0} a_k^0|u_k\rangle\langle u_k|,$$
$$V_1A^{11}V_1^\dagger = \sum_{k=1}^{n_0} a_k^1|v_k\rangle\langle v_k| + \sum_{k=n_0+1}^{n_1} a_k^1|\tilde{v}_k\rangle\langle\tilde{v}_k|$$

for some a_k^0, a_k^1, and $|\tilde{v}_k\rangle$. Note that we can choose $|\tilde{v}_k\rangle$ such that $\forall k, k', k \neq k' : |\tilde{v}_k\rangle\langle\tilde{v}_{k'}| = 0$ and $\forall k, l : |u_k\rangle\langle\tilde{v}_l| = 0$. We can now express $VP_{01}V^\dagger$ as

$$VP_{01}V^\dagger = \sum_{k=1}^{n_0} \big[a_k^0|u_k\rangle\langle u_k| + d_k(|u_k\rangle\langle v_k| + |v_k\rangle\langle u_k|) + a_k^1|v_k\rangle\langle v_k|\big] + \sum_{k=n_0+1}^{n_1} a_k^1|\tilde{v}_k\rangle\langle\tilde{v}_k|.$$

It is now clear that we can choose all $\mathcal{H}_k = \mathrm{span}\{|u_k\rangle, |v_k\rangle\}$, and $\mathcal{H}_{k'} = \mathrm{span}\{|\tilde{v}_{k'}\rangle\}$ which are orthogonal and together add up to $\mathcal{H}$.

In the case that all the d_k are not different, there is some freedom left in choosing $|u_k\rangle$ and $|v_k\rangle$ that still allows us to make $V_0A^{00}V_0^\dagger$ and $V_1A^{11}V_1^\dagger$ diagonal so that the rest of the proof follows in the same way. ∎

In particular, the previous lemma implies that the posterior states corresponding to strings x for which $f(x) = 0$ are orthogonal to those corresponding to strings x for which $f(x) = 1$, which is expressed in the following lemma.

Lemma VI.2: Suppose one performs the measurement given by $\{\Pi_i : i \in [m]\}$. If the outcome of the measurement is i and the state was $U_b|x\rangle$, then the posterior state is

$$|x, i, b\rangle = \frac{\Pi_i U_b|x\rangle}{\sqrt{\langle x|U_b^\dagger \Pi_i U_b|x\rangle}}.$$

The posterior states satisfy

$$\forall x \in f^{-1}(0), x' \in f^{-1}(1), i \in [m] : \langle x, i, b|x', i, b\rangle = 0.$$

Proof: The proof follows straightforwardly from that fact that the Π_i commute with both P_{00} and P_{01} (which follows from Lemma VI.1). ∎

Now we are ready to prove the main theorem of this section.

Theorem VI.3: Let $|\mathcal{Y}| = |\mathcal{B}| = 2$, then there exists a strategy for Bob such that he succeeds at PI_1-$\mathrm{STAR}(\mathcal{E})$ with probability $p = 1$, for any function f and prior P_X on $\mathcal{X}$ which is uniform on the pre-images $f^{-1}(y)$.

Proof: The strategy that Bob uses is the following.

- Bob performs the measurement given by $\{\Pi_i : i \in [m]\}$.
- He will obtain an outcome $i \in [m]$ and store the posterior state which is supported in the at most two-dimensional subspace $\mathcal{H}_i$.
- After the basis $b \in \{0, 1\}$ is announced, he measures $\{P_{0b}, P_{1b}\}$ and reports the outcome of this measurement.

By Lemma VI.2 this performs with success probability 1. ∎

Our result also gives us a better lower bound for all Boolean functions than what we had previously obtained in Section IV–A. Instead of storing the qubit, Bob now measures it immediately along the lines of Lemma III.1. It is easy to see that for one qubit the worst case posterior states to distinguish are in fact those in Lemma III.1.

Corollary VI.4: Let $|\mathcal{Y}| = |\mathcal{B}| = 2$, then Bob succeeds at PI_0-$\mathrm{STAR}(\mathcal{E})$ with probability at least $p \geq (1 + 1/\sqrt{2})/2$. □

In particular, our result implies that for the task of constructing Rabin-OT in [13] it is essential for Alice to choose a random function f from a larger set, which is initially unknown to Bob.

As a final remark, note that in this result, because we succeed with probability 1, the prior distributions do not play any role. Likewise, it is not actually important that the states ρ_{yb} are proportional to projectors: all that is needed in the most general formulation of the discrimination problem at the beginning is that for both $b \in \{0, 1\}$, the states ρ_{0b} and ρ_{1b} are orthogonal.

B. Using Three Bases

We have just shown that Bob can compute any Boolean function perfectly when two bases are used. However, we now show

that for any balanced Boolean function there exist three bases, such that Bob needs to store *all* qubits, in order to compute the function perfectly. The idea behind our proof is that for a particular choice of three bases, any measurement operator that satisfies the conditions set out in Lemma V.1 must be proportional to the identity. This means that we cannot reduce the number of qubits to be stored by a measurement and must keep everything. First, we prove the following lemma which we will need in our main proof.

Lemma VI.5: Let M be a self-adjoint matrix which is diagonal in two MUBs, then M must be proportional to the identity.

Proof: Let $|x\rangle, |u_x\rangle$ with $x \in \{1, \ldots, d\}$ be the basis vectors of the two MUBs and let m_x be the eigenvalue corresponding to $|x\rangle$ and $|u_x\rangle$, then we can write

$$M = \sum_{x=1}^{d} m_x |x\rangle\langle x| = \sum_{x'=1}^{d} m_{x'} |u_{x'}\rangle\langle u_{x'}|.$$

From the previous equation, it follows that

$$\langle x|M|x\rangle = m_x = \sum_{x'=1}^{d} m_{x'} |\langle u_{x'}|x\rangle|^2 = \frac{1}{d}\mathrm{Tr}M$$

which implies the desired result. ∎

We are now ready to prove the main result of this section.

Theorem VI.6: Let $|\mathcal{Y}| = 2$ and $|\mathcal{B}| = 3$, then for any balanced function f and prior P_X on $\mathcal{X}$ which is uniform on the pre-images $f^{-1}(y)$, there exist three bases such that Bob succeeds at PI_q-$\mathrm{STAR}(\mathcal{E})$ with probability $p = 1$ if and only if $q = \log d$.

Proof: Let $P_{00} = \sum_{x \in f^{-1}(0)} |x\rangle\langle x|$, $P_{01} = U_1 P_{00} U_1^\dagger$ and $P_{02} = U_2 P_{00} U_2^\dagger$. Also, let $s : f^{-1}(0) \to f^{-1}(1)$ be a bijective map, and let $s_x = s(x)$. By a reordering of the basis P_{00}, U_1 and U_2 can be written as

$$P_{00} = \begin{bmatrix} \mathbb{I} & 0 \\ 0 & 0 \end{bmatrix}$$
$$U_1 = \begin{bmatrix} U_1^{00} & U_1^{01} \\ U_1^{10} & U_1^{11} \end{bmatrix}$$
$$U_2 = \begin{bmatrix} U_2^{00} & U_2^{01} \\ U_2^{10} & U_2^{11} \end{bmatrix}$$

where all the blocks are of size $(d/2) \times (d/2)$. P_{01} and P_{02} then take the following form:

$$P_{01} = \begin{bmatrix} U_1^{00}U_1^{00\dagger} & U_1^{00}U_1^{10\dagger} \\ (U_1^{00}U_1^{10\dagger})^\dagger & U_1^{10}U_1^{10\dagger} \end{bmatrix}$$
$$P_{02} = \begin{bmatrix} U_2^{00}U_2^{00\dagger} & U_2^{00}U_2^{10\dagger} \\ (U_2^{00}U_2^{10\dagger})^\dagger & U_2^{10}U_2^{10\dagger} \end{bmatrix}.$$

It follows from Lemma V.1, that we only need to prove that $[M, P_{00}] = [M, P_{01}] = [M, P_{02}] = 0$ implies that M must be proportional to the identity. Write

$$M = \begin{bmatrix} M^{00} & M^{01} \\ (M^{01})^\dagger & M^{11} \end{bmatrix}.$$

Commutation with P_{00} implies $M^{01} = 0$. Commutation with P_{01} and P_{02} implies

$$[M^{00}, U_1^{00}U_1^{00\dagger}] = [M^{00}, U_2^{00}U_2^{00\dagger}] = 0 \quad (19)$$
$$[M^{11}, U_1^{10}U_1^{10\dagger}] = [M^{11}, U_2^{10}U_2^{10\dagger}] = 0 \quad (20)$$
$$M^{00}(U_1^{00}U_1^{10\dagger}) = (U_1^{00}U_1^{10\dagger})M^{11} \quad (21)$$
$$M^{00}(U_2^{00}U_2^{10\dagger}) = (U_2^{00}U_2^{10\dagger})M^{11}. \quad (22)$$

We choose U_1 and U_2 in the following way:

$$U_1 = \sum_{x \in f^{-1}(0)} \Big[a_x(|x\rangle\langle x| + |s_x\rangle\langle s_x|) + \sqrt{1-a_x^2}(|x\rangle\langle s_x| - |s_x\rangle\langle x|)\Big]$$
$$U_2 = \sum_{x \in f^{-1}(0)} \Big[a_x(|u_x\rangle\langle u_x| + |v_x\rangle\langle v_x|) + \sqrt{1-a_x^2}(|u_x\rangle\langle v_x| - |v_x\rangle\langle u_x|)\Big]$$

with $a_x \in [0,1]$, satisfying $a_x = a_{x'}$ if and only if $x = x'$. Furthermore, choose $|u_x\rangle$ and $|v_x\rangle$ such that

$$\forall x, x' \in f^{-1}(0), \quad \langle x|v_{x'}\rangle = \langle s_x|u_{x'}\rangle = 0$$
$$|\langle x|u_{x'}\rangle|^2 = |\langle s_x|v_{x'}\rangle|^2 = 2/d.$$

With this choice for U_1 and U_2 we have that

$$U_1^{00}U_1^{00\dagger} = \sum_{x \in f^{-1}(0)} a_x^2 |x\rangle\langle x|$$
$$U_2^{00}U_2^{00\dagger} = \sum_{x \in f^{-1}(0)} a_x^2 |u_x\rangle\langle u_x|$$

i.e., $\{|x\rangle\}$ and $\{|u_x\rangle\}$ form an eigenbasis for $U_1^{00}U_1^{00\dagger}$ and $U_2^{00}U_2^{00\dagger}$, respectively. Furthermore, since all the a_x^2 are different, the eigenbases are unique. Now, using (19), we see that M^{00} must commute with both $U_1^{00}U_1^{00\dagger}$ and $U_2^{00}U_2^{00\dagger}$, and since their eigenbases are unique, it must be true that M^{00} is diagonal in both $\{|x\rangle\}$ and $\{|u_x\rangle\}$. Using the result of Lemma VI.5 it follows that $M^{00} = m_0 \mathbb{I}_{d/2}$. In exactly the same way we can prove that $M^{11} = m_1 \mathbb{I}_{d/2}$ using (20). It remains to prove that $m_0 = m_1$, which follows directly from either (21) or (22). ∎

From our proof it is clear how to construct appropriate U_1 and U_2. For example: Let P_{00} be as defined above, and choose vectors of the form $|x\rangle = |0\rangle|\hat{x}\rangle$ and $|s_x\rangle = |1\rangle|\hat{x}\rangle$ where $\hat{x} \in \{0,1\}^{n-1}$ to construct U_1. For U_2 pick $|u_x\rangle = |0\rangle H^{\otimes n-1}|\hat{x}\rangle$ and analogously $|v_x\rangle = |1\rangle H^{\otimes n-1}|\hat{x}\rangle$.

Note that whereas we know that for such unitaries Bob must store all qubits in order to compute the value of the function perfectly, it remains unclear how close he can come to computing the function perfectly. In particular, he can always choose two of the three bases, and employ the strategy outlined in the previous section: he stores the one qubit that allows him to succeed with probability 1. If he gets the third basis then he just flips a coin. In this case, he is correct with probability $2/3 + 1/(3 \cdot 2) = 5/6$ for a balanced function and a uniform prior.

VII. Conclusion and Open Questions

We have introduced a new state discrimination problem, motivated by cryptography: discrimination with extra information about the state after the measurement, or, more generally, after a quantum memory bound applies. We have left most general questions open, but we found fairly complete results in the case of guessing $y = f(x)$ with mutually unbiased encodings.

We have shown that storing just a single qubit allows Bob to succeed at PI-STAR perfectly for *any* Boolean function and any two bases. On the contrary, we showed how to construct *three* bases such that Bob needs to store *all* qubits in order to compute the function perfectly.

We have also given an explicit strategy for two functions, namely the AND and the XOR. More generally, it would be interesting to find out, how many qubits Bob needs to store to compute $f(x)$ perfectly for any function $f: \mathcal{X} \to \mathcal{Y}$ in terms of the number of outputs $|\mathcal{Y}|$ and the number of bases $|\mathcal{B}|$. It should be clear that the algebraic techniques of Section V allow us to answer these questions for any given function in principle. However, so far, we have not been able to obtain general structures for wider classes of functions.

Our results imply that in existing protocols in the bounded quantum storage model [13] we cannot restrict ourselves to a single fixed function f. However, a great challenge arises in considering more than one function, where f is also announced after the memory bound applies [13].

In general, it is an interesting problem to consider when post-measurement information is useful and how large the advantage can be for Bob. In the important case of two MUBs and balanced functions, we have shown (Theorem III.3 and Corollary IV.4) that there exists a clear separation between the case where Bob gets the post-measurement information (PI-STAR) and when he does not (STAR). Namely, for any such function, Bob's optimal success probability is never larger than $(1 + 1/\sqrt{2})/2 \approx 0.853$ for STAR and always at least as large as the same number for PI-STAR.

In some cases the gap between STAR and PI-STAR can be more dramatic. The XOR function on strings of even length with two MUBs is one of these cases. We have shown that in this case, the advantage can be maximal. Namely, *without* the extra information Bob can never do better than guessing the basis, *with* it however, he can compute the value of the function perfectly. This contrasts with the XOR function on strings of odd length, where the optimal success probabilities of STAR and PI-STAR are both $(1 + 1/\sqrt{2})/2$ and the post-measurement information is completely useless for Bob. It would be interesting to see, how large the gap between STAR and PI-STAR can be for any function $f: \{0,1\}^n \to \{0,1\}^k$ where $k > 2$. It would also be nice to show a general lower bound for nonbalanced functions or a nonuniform prior. As the example for three bases showed, the uniform prior is not necessarily the one that leads to the largest gap, and thus the prior can play an important role. Another generalization would be to consider functions of the form $f: [d]^n \to [d]^k$.

We close by pointing out a potentially interesting connection to the problem of information locking with MUBs [15] and random bases [21]. There, the objective is not so much to obtain an accurate guess of the value $y = f(x)$, as to maximize the (classical) mutual information at the end. In locking, we distinguish measurement with basis information, analogous to our PI_q-STAR with $q = n$, and without (or rather only after the measurement), corresponding to PI_0-STAR. From a classical perspective, it is surprising that the difference in attainable accessible information between $q = n$ and $q = 0$ can be much larger than the information contained in a message specifying which basis was used. In our scenario, we are not interested in locking a string x, but in locking $f(x)$ for a fixed function f. The strength of the observed locking effect depends on the ratio of the number of values f can take and the number of bases used. The dependence on the number of bases carries over to information locking [21], but see the cautionary tale of [3]. It would be interesting to generalize information locking to intermediate values of q, but it seems that we first need to understand the intricate conditions the *bases* have to meet to ensure locking in the first place.

Appendix
Optimal Measurement for the AND Function

For the interested reader, we present the line of thought of deriving the optimal measurement of computing the AND function, when we are allowed to use post-measurement information. Here, Bob is not allowed to store any qubits.

Supported by numerical calculations we construct the following measurement:

$$M_{00} = \Pi_\perp$$
$$M_{o_+o_\times} = \lambda_{o_+o_\times} |\psi_{o_+o_\times}\rangle\langle\psi_{o_+o_\times}|$$

for some $|\psi_{o_+o_\times}\rangle \in \mathcal{H}_2$, for $o_+o_\times \neq 00$ chosen later. Since $\Pi_{||}|\psi_{o_+o_\times}\rangle = |\psi_{o_+o_\times}\rangle$, we can express Bob's probability of success using such a measurement as

$$p = \frac{1}{4}\left[\mathrm{Tr}[b_{00}\Pi_\perp] + \sum_{o_+o_\times \in \{0,1\}, o_+o_\times \neq 00} \right. \tag{23}$$

$$\left. \mathrm{Tr}[\Pi_{||} b_{o_+o_\times} \Pi_{||} M_{o_+o_\times}]\right]. \tag{24}$$

When we project b_{01}, b_{10}, and b_{11} onto $\mathcal{H}_2$ we obtain

$$\begin{aligned}
\Pi_{||}\rho_{0+}\Pi_{||} &= \frac{|c_0\rangle\langle c_0|}{2^n - 1} \\
\Pi_{||}\rho_{0\times}\Pi_{||} &= \frac{|h_0\rangle\langle h_0|}{2^n - 1} \\
\Pi_{||}\rho_{1+}\Pi_{||} &= \rho_{1+} = |c_1\rangle\langle c_1| \\
\Pi_{||}\rho_{1\times}\Pi_{||} &= \rho_{1\times} = |h_1\rangle\langle h_1|.
\end{aligned} \tag{25}$$

Substituting into (23) we then get

$$\begin{aligned}
p = \frac{1}{4}\Bigg[2\frac{2^n - 2}{2^n - 1} &+ \frac{\langle c_0|M_{01}|c_0\rangle}{2^n - 1} + \langle h_1|M_{01}|h_1\rangle \\
&+ \frac{\langle h_0|M_{10}|h_0\rangle}{2^n - 1} + \langle c_1|M_{10}|c_1\rangle + \langle c_1|M_{11}|c_1\rangle \\
&+ \langle h_1|M_{11}|h_1\rangle\Bigg].
\end{aligned}$$

Now using that $M_{11} = \Pi_{||} - M_{01} - M_{10}$ we get

$$p = \frac{1}{4}\left[2\frac{2^n-2}{2^n-1} + \frac{\langle c_0|M_{01}|c_0\rangle}{2^n-1} - \langle c_1|M_{01}|c_1\rangle + \frac{\langle h_0|M_{10}|h_0\rangle}{2^n-1} - \langle h_1|M_{10}|h_1\rangle + 2\right].$$

We now show how to choose $|\psi_{o_+o_\times}\rangle$. We take $\lambda = \lambda_{01} = \lambda_{10}$. Then s_1 and s_2, the only possible nonzero eigenvalues of

$$M_{01} + M_{10} = \lambda(|\psi_{01}\rangle\langle\psi_{01}| + |\psi_{10}\rangle\langle\psi_{10}|)$$

satisfy

$$s_1 + s_2 = 2\lambda$$
$$s_1^2 + s_2^2 = 2\lambda^2(1 + |\langle\psi_{01}\psi_{10}\rangle|^2).$$

From these two equations one gets

$$s_1 = \lambda(1+\eta)$$
$$s_2 = \lambda(1-\eta)$$

where $\eta e^{i\phi} = \langle\psi_{10}|\langle\psi_{01}\rangle$.[1] Recall that we need to have $M_{01} + M_{10} + M_{11} = \Pi_{||}$, which is equal to the identity on $\mathcal{H}_2$. We therefore want one of the eigenvalues of $M_{01}+M_{10}$ to be 1 and the other one smaller than 1. So we must choose $\lambda = (1+\eta)^{-1}$. We then also need

$$\lambda_{11} = \frac{2\eta}{1+\eta}$$
$$|\psi_{11}\rangle = \frac{|\psi_{01}\rangle - e^{i\phi}|\psi_{10}\rangle}{\sqrt{2(1-\eta)}}.$$

We then take, supported by the symmetry of the problem

$$|\psi_{01}\rangle = \alpha|c_0\rangle + \beta|c_1\rangle$$
$$|\psi_{10}\rangle = \alpha|h_0\rangle + \beta|h_1\rangle$$

with α and β real and satisfying $\alpha^2+\beta^2 = 1$. We have now that

$$\langle\psi_{10}|\psi_{01}\rangle = (-1)^{n+1}\frac{\alpha^2-\beta^2}{2^{n/2}} + 2\alpha\beta\sqrt{1-\frac{1}{2^n}}.$$

Now p becomes

$$p = \frac{1}{2}\left[\frac{2^n-2}{2^n-1} + \frac{1}{1+\eta}\left(\frac{\alpha^2}{2^n-1} - \beta^2\right) + 1\right]$$
$$= \frac{1}{2}\left[\frac{2^n-2}{2^n-1} + \frac{1}{1+\eta}\frac{1-2^n\beta^2}{2^n-1} + 1\right]$$

and

$$\eta = \left|\frac{\alpha^2-\beta^2+(-1)^{n+1}2\alpha\beta\sqrt{2^n-1}}{2^{n/2}}\right|$$
$$= \left|\frac{1-2\beta^2+(-1)^{n+1}2\beta\sqrt{1-\beta^2}\sqrt{2^n-1}}{2^{n/2}}\right|$$

[1]In the present context $e^{i\phi} = \pm 1$.

where w.l.o.g. we have chosen α to be positive. We want η to be small. It is easy to see that we would like to take $\beta = (-1)^n|\beta'|$, for some real β'. A simple calculation shows that then to minimize η we should choose

$$|\beta'| = \frac{1}{\sqrt{2^{2n} + 2^{\frac{3}{2}n+1} - 2^{\frac{n}{2}+1}}}.$$

ACKNOWLEDGMENT

The authors wish to thank Harry Buhrman for his persistent interest in the present investigation, various discussions, and his suggestion that our problem also has applications to communication complexity. Thanks also to Ronald de Wolf for helpful comments on an earlier version of the manuscript.

REFERENCES

[1] E. Andersson, S. Barnett, C. Gilson, and K. Hunter, "Minimum-error discrimination between three mirror-symmetric states," *Phys. Rev. A*, vol. 65, p. 052308, 2002.

[2] Y. Aharonov and B.-G. Englert, "The mean king's problem: Prime degrees of freedom," *Phys. Lett. A*, vol. 284, pp. 1–5, 2001.

[3] M. Ballester and S. Wehner, "Entropic uncertainty relations and locking: tight bounds for mutually unbiased bases," *Phys. Rev. A*, vol. 75, p. 022319, 2007.

[4] M. Ban, K. Kurokawa, R. Momose, and O. Hirota, "Optimum measurements for discrimination among symmetric quantum states and parameter estimation," *Int. J. Theor. Phys.*, vol. 36, no. 6, pp. 1269–1288, 1997.

[5] S. Bandyopadhyay, P. O. Boykin, V. P. Roychowdhury, and F. Vatan, "A new proof for the existence of mutually unbiased bases," *Algorithmica*, vol. 34, no. 4, pp. 512–528, 2002.

[6] S. M. Barnett, "Minimum-error discrimination between multiply symmetric states," *Phys. Rev. A*, vol. 64, p. 030303, 2001.

[7] J. Bergou, U. Herzog, and M. Hillery, "Quantum state filtering and discrimination between sets of boolean functions," *Phys. Rev. Lett.*, vol. 90, p. 257901, 2003.

[8] J. Bergou, U. Herzog, and M. Hillery, "Discrimination of quantum states," in *Quantum State Estimation*, M. Paris and J. Rehacek, Eds. Berlin, Germany: Springer-Verlag, 2004, vol. 3, pp. 417–465.

[9] J. Bergou, U. Herzog, and M. Hillery, "Optimal unambiguous filtering of a quantum state: An instance in mixed state discrimination," *Phys. Rev. A*, vol. 71, p. 042314, 2005.

[10] J. Bergou and M. Hillery, "Quantum-state filtering applied to the discrimination of boolean functions," *Phys. Rev. A*, vol. 72, p. 012302, 2005.

[11] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[12] O. Bratteli and D. W. Robinson, "Operator algebras and quantum statistical mechanics. 1. C^*- and W^*-algebras, symmetry groups, decomposition of states," in *Texts and Monographs in Physics*, 2nd ed. Berlin, Germany: Springer-Verlag, 1987.

[13] I. Damgaard, S. Fehr, L. Salvail, and C. Schaffner, "Cryptography in the bounded quantum-storage model," in *Proc. 46th IEEE Conf. Foundations of Computer Science (FOCS)*, Pittsburgh, PA, Oct. 2005, pp. 449–458.

[14] G. M. D'Ariano, D. Kretschmann, D. Schlingemann, and R. F. Werner, Quantum Bit Commitment: The Possible and the Impossible, 2006,, quant-ph/0605224.

[15] D. P. DiVincenzo, M. Horodecki, D. W. Leung, J. A. Smolin, and B. W. Terhal, "Locking classical correlation in quantum states," *Phys. Rev. Lett.*, vol. 92, p. 067902, 2004.

[16] Y. Eldar, "A semidefinite programming approach to optimal unambiguous discrimination of quantum states," *IEEE Trans. Inf. Theory*, vol. 49, no. 2, pp. 446–456, Feb. 2003.

[17] Y. Eldar and G. Forney, "On quantum detection and the square-root measurement," *IEEE Trans. Inf. Theory*, vol. 47, no. 3, pp. 858–872, Mar. 2001.

[18] Y. Eldar, A. Megretski, and G. Verghese, "Designing optimal quantum detectors via semidefinite programming," *IEEE Trans. Inf. Theory*, vol. 49, no. 4, pp. 1007–1012, Apr. 2003.
[19] Y. Eldar, A. Megretski, and G. Verghese, "Optimal detection of symmetric mixed quantum states," *IEEE Trans. Inf. Theory*, vol. 50, no. 6, pp. 1198–1207, Jun. 2004.
[20] P. Hausladen and W. K. Wootters, "A pretty good measurement for distinguishing quantum states," *J. Mod. Opt.*, vol. 41, pp. 2385–2390, 1994.
[21] P. Hayden, D. Leung, P. Shor, and A. Winter, "Randomizing quantum states: Constructions and applications," *Commun. Math. Phys.*, vol. 250, no. 2, pp. 371–391, 2004.
[22] C. W. Helstrom, "Quantum detection and estimation theory," *J. Stat. Phys.*, vol. 1, no. 2, pp. 231–252, 1969.
[23] A. S. Holevo, "Statistical decision theory for quantum systems," *J. Multivariate Anal.*, vol. 3, no. 337, 1973.
[24] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1985.
[25] K. Hunter, "Measurement does not always aid state discrimination," *Phys. Rev. A*, vol. 68, p. 012306, 2003.
[26] K. Hunter, "Results in optimal discrimination," in *Proc. QCMC*, Glasgow, U.K., Jul. 2004, pp. 83–86.
[27] A. Klappenecker and M. Rötteler, Frakfamily New Tales of the Mean King 2005, quant-ph/0502138.
[28] M. Koashi and N. Imoto, "Operations that do not disturb partially known quantum states," *Phys. Rev. A*, vol. 66, p. 022318, 2002.
[29] H.-K. Lo and H. F. Chau, "Is quantum bit commitment really possible?," *Phys. Rev. Lett.*, vol. 78, pp. 3410–3413, 1997.
[30] L. Masanes, "Asymptotic violation of Bell inequalities and distillability," *Phys. Rev. Lett.*, vol. 97, p. 050503, 2006.
[31] D. Mayers, The Trouble with Quantum Bit Commitment 1996, quant-ph/9603015.
[32] C. Mochon, A family of Generalized 'Pretty Good' Measurements and the Minimal-Error Pure-State Discrimination Problems for which they are Optimal quant-ph/0506061.
[33] M. Takesaki, *Theory of Operator Algebras. I.* Berlin, Germany: Springer-Verlag, 1979.
[34] M. Wang and F. Yan, Conclusive Quantum State Classification quant-ph/0605127.
[35] P. Wocjan and T. Beth, "New construction of mutually unbiased bases in square dimensions," *Quantum Inf. Comput.*, vol. 5, no. 2, pp. 129–158, 2005.
[36] W. K. Wootters and B. Fields, "Optimal state-determination by mutually unbiased measurements," *Ann. Phys.*, vol. 191, no. 368, 1989.
[37] H. P. Yuen, R. S. Kennedy, and M. Lax, "Optimum testing of multiple hypotheses in quantum detection theory," *IEEE Trans. Inf. Theory*, vol. IT-21, no. 2, pp. 125–134, Mar. 1975.